



# **iJRASET**

International Journal For Research in  
Applied Science and Engineering Technology



---

# **INTERNATIONAL JOURNAL FOR RESEARCH**

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 10    **Issue:** XII    **Month of publication:** December 2022

**DOI:** <https://doi.org/10.22214/ijraset.2022.47981>

**[www.ijraset.com](http://www.ijraset.com)**

**Call:** ☎ 08813907089

**E-mail ID:** [ijraset@gmail.com](mailto:ijraset@gmail.com)

# Four-Character Captcha Recognition using Deep Learning Techniques

Aitha Sahith<sup>1</sup>, P Swethan<sup>2</sup>, Balle Naresh<sup>3</sup>

<sup>1, 2, 3</sup>Sreenidhi Institute of Science and Technology Yamnampet, Ghatkesar, Hyderabad, Telangana-501301

**Abstract:** *The Completely Automated Public Turing Test to Tell Computers and Humans Apart (CAPTCHA) is a crucial human-machine distinguishing technique for websites to fend against automated attacks from malicious software. Studies on CAPTCHA recognition have the potential to uncover security flaws, improve CAPTCHA technology, and even advance handwriting and license plate recognition. However, the whole objective of the CAPTCHAs may be bypassed using the principles of deep learning and computer vision. Convolutional neural networks (CNN) may be used to automatically pass this exam. A CNN is a deep learning algorithm that uses an image as input before giving distinct characteristics in the picture a value that further aids in differentiating one feature from another. Its major objective is to convert the pictures into a more simple form while retaining elements that are crucial for obtaining an optimum forecast.*

**Keywords:** CAPTCHA, human, machine, CNN, algorithm, deep learning, feature extraction.

## I. INTRODUCTION

A test to distinguish between a person and a machine is called the Completely Automated Public Turing Test to Tell Computers and Humans Apart, or CAPTCHA. [1] The vast majority of Web apps employ this test. A text-based, visual-based, or auditory challenge-response system is the foundation of the CAPTCHA reverse Turing test. Many implicit human interactions with services on the Internet use CAPTCHA tactics to confirm that a user is a person and not a Web-bot. Chat rooms, search engines, password systems, online polls, e-mail services for account registrations, prevention of sending and receiving spam, blogs, messaging services, free content downloading services, and phishing attack detection are among the web services and applications that use CAPTCHA methods [7]. The application's security is improved with CAPTCHA. Initial CAPTCHAs are picture-based, requiring the user to recognise the item in the image in order to be confirmed. After correctly entering the name of the thing in the image, the user is verified. Google has suggested a CAPTCHA technique that requires the user to turn pictures that have been randomly rotated so that they are upright. [6]. There are other CAPTCHAs that employ audio. It is essential in many security applications since it prevents bots from automatically registering users [4]. When creating an audio CAPTCHA, a random series of basic words or numbers is recorded, combined, and then noise and disruption are added. This is just a warped visual that a computer has a very difficult time understanding, in contrast to a person. Text-based CAPTCHAs take the form of a picture with a cryptic text string that must be recognised and entered by the user in a text box that is supplied next to the CAPTCHA image on the website. The CAPTCHA picture is of poor quality, with many types of noise and significant deterioration [5]. This prevents automated systems from using sensitive or restricted data without authorization. Even though robots are unable to detect the letters in each CAPTCHA picture, technologies like deep learning may be used to train them to do so in a manner similar to how a normal person would.

Deep learning is a branch of machine learning that is based on a kind of architecture known as a neural network. In this case, the computer duplicates the analogy of human learning. To comprehend and evaluate the data supplied to it, the machine mimics the functioning of neuron cells in the human brain. Problem statements that incorporate multi-variate, high-dimensional data, such as photos and movies, often employ deep learning. Deep learning models' strong design allows them to manage vast amounts of complicated data.

The technique of converting a randomly generated CAPTCHA's input picture into digital or raw text is known as CAPTCHA recognition. To do this, create a deep learning model that is trained to recognise the characters in the provided CAPTCHA picture. Convolutional neural networks (CNN) are the foundation of this approach. An application's vulnerability may be ascertained by looking at the CAPTCHA it uses.

Systems that recognise CAPTCHA may be used to assess a website's or application's security levels. The model that can decipher a CAPTCHA can also extract text from an image more effectively than traditional systems, and this model may be improved to identify intricate handwriting.

## II. RELATED WORKS

In a segmented picture, traditional approaches often detect a single number or character region and then identify a single character. Microsoft CAPTCHA was successfully segmented and identified by a number of classifiers [8], but the recognition rate was just 60%. a technique for recognition using LSTM (Long Short-Term Memory Network) recurrent neural networks (RNNs) [9]. English letters were encoded using adaptive binary arithmetic coding by Lingyun Xiang et al.

There was a good outsourcing word segmentation technique [10] that conserved storage space. A self-organizing incremental neural network based on the PNN-SOINN-RBF described in [11]. On the verification set using both offline and online models, the overall prediction accuracy of single characters was 72.75% and 50.25%, respectively. Using the approach of character segmentation coupled with contour difference projection and a water droplet algorithm, a three-color denoising method based on RGB has been discussed [12]. With regard to the CAPTCHA's background noise and character distortion adhesion, our technique had great recognition results.

Because it may compromise the identification accuracy of these systems, the character recognition module is also regarded as a vital part of segmentation-based CAPTCHA recognition systems. By comparing distinct characters with template characters using character coefficient values, the template matching [13,14,15] technique is used to identify characters. If more sophisticated techniques are not used, errors resulting from character similarity are seen as a weak spot in recognition by template matching. A CAPTCHA character recognition technique called selective learning confusion class was presented in [16]. To identify CAPTCHA characters, SLCC employs a two-stage Deep Convolutional Neural Network (DCNN) frame. The characters are first classified using an all-class DCNN classifier. Then, confusion class subsets are created using a confusion relation matrix and a set partition procedure. This CAPTCHA character recognition approach provides excellent character identification accuracy, particularly for characters in the confusion class; nevertheless, by creating a new DCNN for each subset of the confusion class, the total system storage capacity may expand significantly.

Major organizations have already committed large resources to overcoming CAPTCHAs in order to evaluate the benefits and drawbacks of these data strategies since CAPTCHAs are actively employed by many websites to safeguard traffic. For instance, the Street View team at Google was able to solve the CAPTCHA challenge with 99.8% [17] accuracy using their algorithms for identifying signs in photographs.

Although there are many other types of CAPTCHAs, text-based CAPTCHAs are often employed despite not being the most secure choice since they are affordable, practical, and user-friendly. Text-based CAPTCHAs need to be improved since they are less secure than planned and more open to assaults. By identifying their flaws, it is possible to strengthen the security of text-based CAPTCHAs by discovering more effective and accurate solutions. Overall, CNN is a reliable and effective technique for identifying CAPTCHAs, and future work may focus on enhancing the security of text-based CAPTCHAs by using it. Our team built three CNN networks [18] that are structurally more efficient than many of the existing methods of high-accuracy CAPTCHA recognition and tested them on three different CAPTCHA datasets to see if they could be used to accurately recognise CAPTCHAs with character adhesions and background noise.

For the purpose of directly identifying CAPTCHAs without segmentation, researchers have lately started to use deep learning-based CAPTCHA identification algorithms. To increase the precision of character identification, the internal organization of this CNN model was created to take into account the association between neighboring characters. The architecture of this model is considerably complicated since each character is represented by a unique collection of convolutional layers.

## III. METHODOLOGY

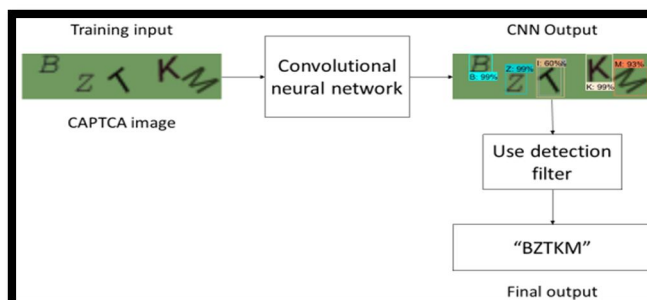


Fig. 1: Workflow



On top of the CNN architecture, a deep learning model is developed. This model of four-character CAPTCHA recognition can identify the characters in a given CAPTCHA. A random combination of 36 alphanumeric characters (26 alphabetic and 10 numeric) is used to create these four-character CAPTCHAs. The model is made using the Tensor Flow framework. To create this model, a Python script is built. 2000 randomly generated pictures from the Captcha package and random module in Python make up the dataset for training. [2] . The picture has noise added to it to make it a little harder to identify. Utilizing a True Type Font (TTF) file with the proper font style for the characters in the CAPTCHA will allow you to get these pictures. The characters found in each picture are used to label each image in the collection. To cut down on computing costs and training time, each picture is turned into a grayscale version. With the help of a preset function from the OpenCV package, contours are created to identify the characters (or features, in this case) in the provided picture. Simple contour definition: A curve connecting all continuous points (along the border) of the same hue or intensity. For form analysis, item identification, and recognition, contours are a valuable tool [3].

#### IV. IMPLEMENTATION ANALYSIS

The dataset is now used to create the preprocessed picture data. To do this, each picture is divided into four pieces, and the sliced images that make up a single character are then distributed among different classes according to the labels that were already defined when the dataset was formed. This preprocessed data is used during model training. The process currently involves building a deep net with seven layers, including two convolutional layers, two max-pooling layers, a flattening layer, and two dense layers.

ReLU is the activation function utilized in the first layer of the network's 32x32x1 input form with 32 neurons. The first convolution layer uses a square matrix with a kernel size of 3 by 3. A softmax activation function is employed for multiclassification in the last layer, which includes 36 neurons. The model undergoes 50 training iterations with a learning rate of 0.001. After a successful training session, the model is projected onto test data, where its performance may be assessed.

Our CAPTCHA issue is special in that we can conduct learning without requiring human assistance. To do this, we only retrain data samples for which the classifier has previously assigned a valid label. We use the above-mentioned uncertainty levels to prevent this: In each learning cycle, we filter the properly identified test samples by prediction uncertainty and only employ the least certain ones. As a consequence, fewer training examples are needed, but as we shall demonstrate in the experiments, the more ambiguous instances are also the most instructive for learning.

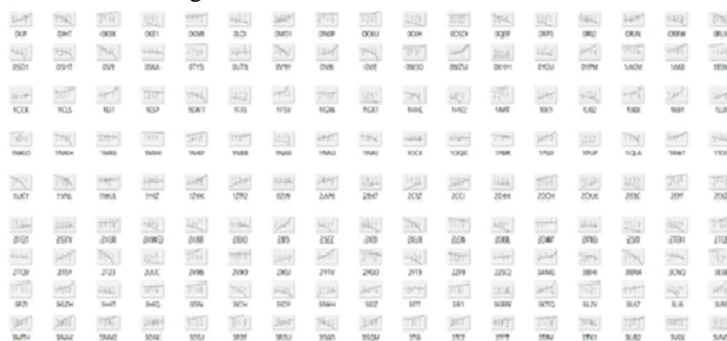


Fig 2 : the collected data set.



Fig 3: the data set After preprocessing

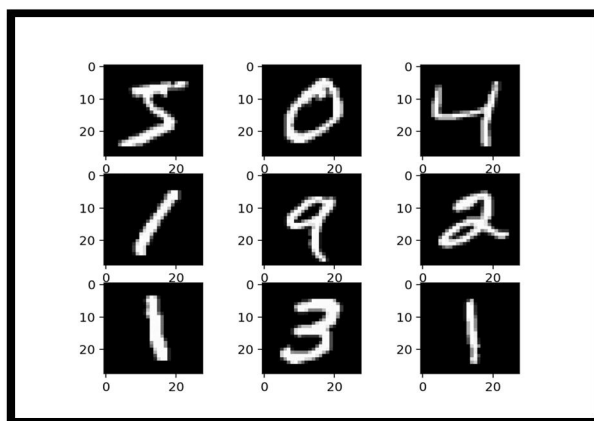


Fig 4. Data after preprocessing

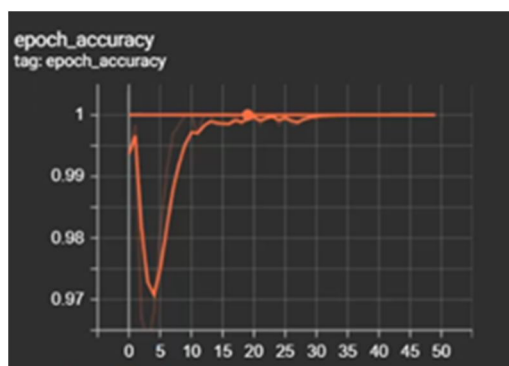


Fig 5: Model accuracy

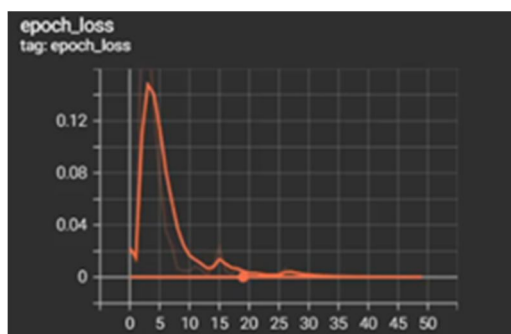


Fig 6: Model loss

Model accuracy: 70%.

## V. CONCLUSION

Deep learning helps computers to derive meaningful links from a plethora of data and make sense of unstructured data. Here, the mathematical algorithms are combined with a lot of data and strong hardware to get qualified information. CAPTCHA was designed to improve the security of the systems but deep learning algorithms defeated its very purpose. Here, we used Convolutional Neural networks for CAPTCHA recognition

## VI. FUTURE ENHANCEMENT

The future scope of this work lies to expand this CAPTCHA recognition system for larger and more noisy CAPTCHA containing all the symbols possible. The enhancement of the project leads to the solution of most interesting challenging problems like Doctor-Handwriting Recognition.

## REFERENCES

- [1] <https://en.wikipedia.org/wiki/CAPTCHA>
- [2] <https://pypi.org/project/claptcha/>
- [3] [https://docs.opencv.org/3.4/d4/d73/tutorial\\_py\\_contours\\_begin.html](https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html)
- [4] <https://ieeexplore.ieee.org/document/9398494>
- [5] <https://arxiv.org/ftp/arxiv/papers/1112/1112.5605.pdf>
- [6] R. Gossweiler, M. Kamvar, and S. Baluja, "Whats Up CAPTCHA? A CAPTCHA Based On Image Orientation", In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, 2009, pages 841-850.
- [7] [https://www.researchgate.net/publication/51967478\\_A\\_Study\\_of\\_CAPTCHAs\\_for\\_Securing\\_Web\\_Services](https://www.researchgate.net/publication/51967478_A_Study_of_CAPTCHAs_for_Securing_Web_Services).
- [8] J. Yan and A. S. E. Ahmad, A low-cost attack on a Microsoft CAPTCHA, Proceedings of the ACM Conference on Computer and Communications Security, (2008), 543–554.
- [9] L. Zhang, S. W. Huang, Z. X. Shi, et al., CAPTCHA recognition method based on LSTM RNN, Pattern Recogn., 1 (2011), 40–47.
- [10] Y. L. Liu, H. Peng and J. Wang, Verifiable diversity ranking search over encrypted outsourced Data, Comput. Mater. Con., 1 (2018), 37–57.
- [11] H. T. Tang, Verification code recognition model and algorithm of self-organizing incremental neural network, MA thesis, Guangdong University of technology, 2016.
- [12] Y. Wang, Y. Q. Xu and Y. B. Peng, Verification code identification of xiaonei network based on KNN technology, Comput. Moder., 2 (2017), 93–97.
- [13] Sakatos, P.; Theerayut, W.; Nuttapol, V.; Surapong, P. Analysis of text-based CAPTCHA images using Template Matching Correlation technique. In Proceedings of the 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE), Chiang Rai, Thailand, 5–8 March 2014; pp. 1–5.
- [14] Wu, C.; On, L.C.; Weng, C.H.; Kuan, T.S.; Ng, K. A Macao license plate recognition system. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, 18–21 August 2005; Volume 7, pp. 4506–4510.
- [15] Baten, R.A.; Omair, Z.; Sikder, U. Bangla license plate reader for metropolitan cities of Bangladesh using template matching. In Proceedings of the 8th International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, 20–22 December 2014; pp. 776–77.
- [16] Chen, J.; Luo, X.; Liu, Y.; Wang, J.; Ma, Y. Selective Learning Confusion Class for Text-Based CAPTCHA Recognition. IEEE Access 2019, 7, 22246–22259.
- [17] Goodfellow, I.J., Bulatov, Y., Ibarz, J. Arnoud, S., Shet, V. (2013). Multi-digit Number Recognition from Street View: Imagery using Deep Convolutional Neural Networks. arxiv preprint.
- [18] Haolin Yang 2020 *J. Phys.: Conf. Ser.* 1693 012040, DOI 10.1088/1742-6596/1693/1/012040.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)