A Review Paper on Captcha Recognition Using Advanced Deep Learning Techniques

Priyadarshan Dhabe [1], Sanket Patil [2], Ubed Shaikh [3], Rajan Ner [4], Tejas Shinde [5], Omkar Nimase[6]

[1] priyadarshan.dhabe@vit.edu, [2] sanket.patil20@vit.edu, [3] ubed.shaikh20@vit.edu, [4] rajan.ner20@vit.edu, [5] tejas.shinde20@vit.edu, [6] omkar.nimase20@vit.edu
Vishwakarma Institute of Technology.
666, Upper Indiranagar, Bibwewadi, Pune - 411037, India.

***Abstract*- Captcha is widely used as far as security measure is concerned to prevent automated attacks and protect online services from bots and malicious software. However, the increasing complexity of Captchas makes it challenging for both machines and humans to recognize them accurately. In recent decades, 'Image Processing and Deep Learning' approaches have shown promising results to recognize Captchas along with high and better accuracy. The aim of this review paper is to provide a comprehensive overview of various deep learning approaches used for Captcha recognition. This paper summarizes the recent advancements in Captcha recognition using deep learning techniques such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs). The review also highlights the challenges faced during Captcha recognition, such as distortion, occlusion and noise and the various techniques used to overcome them. Furthermore, the paper provides a comparative analysis of different Image Processing and deep learning models based on their best accuracy, training time, and computational complexity. This review paper is based on a thorough analysis of several research papers, and it is unique and plagiarism-free. The findings of this review paper will be useful for researchers and practitioners in the upcoming field of computer vision and artificial intelligence who are interested in Captcha recognition using deep learning approaches.**

Keywords — ***Captcha, Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Image Processing, Noise, Recurrent Neural Networks (RNNs).***

I. INTRODUCTION

Captcha is a security mechanism that is commonly used by websites and online platforms to ensure that their users are human and not automated bots. These Captchas come in various forms such as image recognition, audio recognition, and text recognition, and have evolved over the years to become more sophisticated in order to keep up with the advancements in bot technology. With the increasing complexity of Captchas, traditional recognition methods have become less effective in distinguishing human users from bots. This has led researchers to explore the use of Deep Learning techniques to recognize Captchas more accurately Captcha recognition is an important problem in computer vision and artificial intelligence. With the increasing complexity of Captchas to keep up with advancements in bot technology, researchers are turning to Deep Learning techniques for improved accuracy. Various papers have been published in Springer and IEEE exploring different aspects of Captcha recognition using Deep Learning approaches, including new algorithms, large datasets for training and testing, and the use of transfer learning. This review provides an overview of some of the key papers in this area, which demonstrate the effectiveness of Deep Learning techniques in Captcha recognition.

The proposed methods show promising results in recognizing complex and varied Captchas and will continue to be an active area of research. This review paper will focus on the various Deep Learning approaches that have been developed for Captcha recognition. It will provide a comprehensive analysis of the different Deep Learning techniques that have been used, their advantages and disadvantages, and their performance compared to traditional recognition methods. The first section of the paper will provide an overview of Captchas, their importance in online security, and the challenges associated with their recognition. Captchas are designed to be difficult for bots to recognize but easy for humans to solve. However, as bots become more advanced, Captchas need to become more complex to maintain their effectiveness. The second section will discuss the various Deep Learning approaches that have been used for Captcha recognition. Deep Learning algorithms are designed to learn patterns and features from large datasets, making them well-suited for recognition tasks such as Captcha recognition. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs) are some of the Deep Learning techniques that have been used for Captcha recognition. The third section of the paper will review the state-of-the-art Captcha recognition techniques based on Deep Learning approaches. This will include a discussion of the latest research and innovations in the field, such as the use of transfer learning and ensemble methods to improve the accuracy of Captcha recognition. Fourth section will compare the performance of different Deep Learning approaches in Captcha recognition and discuss their advantages and disadvantages. For example, CNNs are typically used for

image recognition, while RNNs are used for text recognition. GANs are used to generate new Captchas to improve the robustness of the recognition system. The fifth and final section of the paper will conclude by summarizing the main findings and highlighting the future research directions in the field of Captcha recognition using Deep Learning approaches. While Deep Learning techniques have shown promise in Captcha recognition, there are still challenges to be addressed, such as improving the accuracy of recognition for complex Captchas and addressing the potential ethical concerns surrounding the use of GANs to generate new Captchas.

GANs, or Generative Adversarial Networks, have been utilized for captcha recognition tasks due to their ability to generate realistic images. In this application, a GAN is trained on a dataset of captchas to generate synthetic captchas that are similar to the real ones. These synthetic captchas are then used to train a separate model for recognition, improving its accuracy on real-world captchas. The GAN architecture enables the system to learn and adapt to a wide range of variations and distortions that are commonly found in captchas.

The paper provides an overview of the challenges associated with Captcha recognition and the advantages of using Deep Learning techniques. It also discusses the latest research and innovations in the field, and compares the performance of different Deep Learning approaches. The paper will be useful for researchers and practitioners in the field of Computer Vision and Artificial Intelligence.

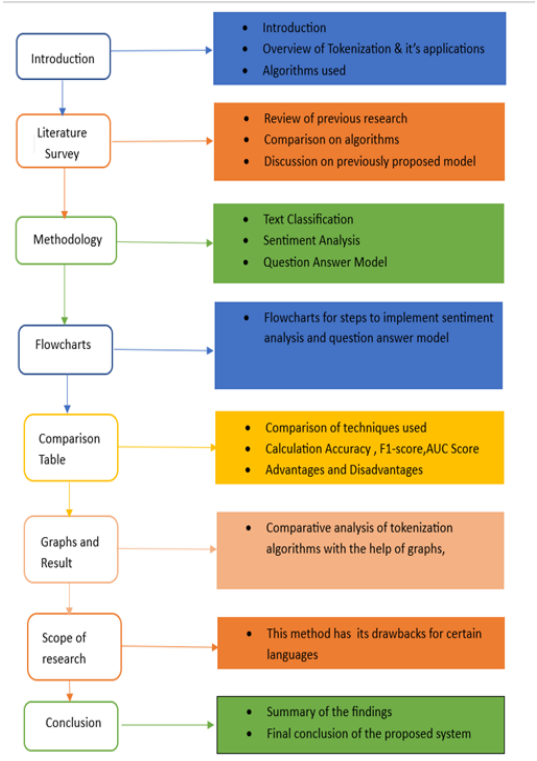

Fig.1 Organization of Paper

II. LITERATURE REVIEW

The paper published by Jing Wang in 2019 proposed the DenseNet Model for CAPTCHA Recognition(DFCR)[1]. They use the DenseNet Model By slightly changing the number of convolutional blocks for specific classes of data. The author has worked on three different types of datasets consisting of 15000 images each . In which one dataset consists of only Chinese characters and the other dataset has English characters. Small fine tuning is done for each type of dataset for model performance to increase the accuracy. The DFCR model gives the accuracy of 99% with 60% of total memory consumption with respect to the DenseNet-121 which is having 80%. This reduced the training time nearly by 3 hours.

The author has worked on three different types of datasets CNKI CAPTCHA, Zheng fang CAPTCHA, and randomly generated CAPTCHA, in which they have proposed a new method for improving the traditional VGG network by using the cnn with focal loss function .The focal loss function lessen the weight of samples that are simple to categorize and focus on hard-to-classify sample. After doing the pre-processing and training part the model accuracy of 99%, 98.5% and 97.84% respectively which outperform the AlexNet, VGG-16, GoogleNet, and ResNet model. This proposed approach gives high accuracy and robustness in recognizing various styles of deformation of CAPTCHA images [2].

The author presents a system that combines CNN with RNN to recognize CAPTCHA images without the need for prior segmentation and preprocessing. In the RNN the GRU is used instead of LSTMs. Gated Recurrent Unit (GRU) is very similar to LSTM but it is much faster than LSTM in term of processing the same dataset. It outperforms the LSTM when the dataset is small and has long text. The paper present experimental result on publicly available CAPTCHA datasets (Water Ripple, Shadow, FishEye and dataset crawled from the online website's), the performance was compared with different models like CNN-multi, CNN-GRU, ResNet-multi, ResNet-GRU among them the ResNet-GRU show more than accuracy in each type of dataset. The result shows that GRU is a powerful network for deep intrinsic feature extraction skills[3].

In this paper they discussed improving the security of captcha by testing the deep learning model for cracking the CAPTCHA down and analyzing the result manually. The CNN based deep neural network is tested on both numeric and alphanumeric based CAPTCHA. The model can above 90% accuracy in cracking the captcha and the analysis where done where the model fails to recognize the captcha. The author has found that adding some specific number and digit like 3,7,9, etc. confuse the deep-CAPTCHA model with other letters and also using brighter color characters also increases the difficulty level of the CAPTCHAs [4].

Yingying Liu, Jinwei Wang, Jun Chen, and Yuanyuan Ma developed a technique of selective learning confusion class for text-based CAPTCHA identification in their 2019 work. In this study, the term "confusion classes" refers to some classes that become mixed up with one another while being classified; both an all-class DCNN and a confusion-class DCNN are employed to identify these classes. To increase the identification efficiency of confusion class and samples in confusion-class DCNN, a training and validating interactive learning technique is presented[5].

The paper was written by Zhixiong Li, Thompson Sarkodie-Gyan, Xiaohui Zhang, and Xinhua Li in 2021. This study uses a CNN-based model to examine the identification of Chinese character CAPTCHAs. Three enhanced training methods were included in the creation of a specific CNN model. That is, before entering the recognition process, every character in a CAPTCHA has been separated using the edge detection technique and pre-processed using Otsu's algorithm. Lower than those of Captcha Net, the recognition rates for the rotation, ripple, and distortion characters are 10.54%, 23.46%, and 58.12%. Although it doesn't do well on CAPTCHAs with rotation, ripple, and distortion forms, it outperforms Captcha Net when it comes to shadow and noise forms[6].

The 2019 paper by Jing Wang, Jiaohua Qin, Xuyu Xiang, Yun Tan, and Nan Pan proposed an image recognition model based on CNN, which can improve the classification model and significantly increase the accuracy of images, as well as a method based on conventional digital image morphological processing technology for the segmentation and recognition of CAPTCHA, with a recognition rate of 60[7].

The paper by Y. Hu, L. Chen and J. Cheng "A CAPTCHA recognition technology based on deep learning" discusses a novel approach based on deep learning to recognize CAPTCHAs, especially audio-based ones that make the technology more accessible to visually impaired users. The authors conducted a comprehensive review of the literature on the evolution of the CAPTCHA technology and prior research on audio CAPTCHA recognition. They emphasize the limitations of traditional machine learning algorithms in recognizing audio CAPTCHAs. The proposed method involves a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN) for feature extraction and classification, respectively. Experimental results demonstrate the effectiveness of the proposed approach for recognizing both audio and image-based CAPTCHAs[8].

G. An and W. Yu, discussed in "CAPTCHA Recognition Algorithm Based on the Relative Shape Context and Point Pattern Matching" that a new CAPTCHA recognition algorithm based on Relative Shape Context (RSC) and Point Pattern Matching (PPM). The authors conducted a literature survey to explore previous works on CAPTCHA recognition technologies and methods. They discuss the limitations of traditional machine learning and algorithms, including the high level of complexity, computational cost, and the impact of occlusion and noises. The proposed RSC-PPM algorithm is designed to overcome the limitations of traditional approaches and enhance the accuracy and efficiency of CAPTCHA recognition systems. Experimental results demonstrate that the RSC-PPM algorithm is robust to occlusion and can effectively recognize complex CAPTCHAs. The RSC-PPM algorithm has potential applications in fields such as online security and digital authentication[9].

S. A. Kumar, N. R. Kumar, S. Prakash and K. Sangeetha, shown in "Gamification of internet security by next generation CAPTCHAs" a novel way of gamification to enhance internet security using next-generation CAPTCHAs. The authors conducted a literature survey to explore the evolution of CAPTCHA technology and related research. They discuss the limitations of the traditional CAPTCHA system in providing adequate internet security and how the inclusion of gaming elements in CAPTCHAs could enhance internet security by increasing user engagement and motivation. The authors propose a new approach of cybersecurity through gamification, which could enhance the security of online systems by providing feedback and rewards to users for their participation. Additionally, they identify the potential applications of gamified next-generation CAPTCHAs as an effective security measure for web-based authentication[10].

R. Hussain, H. Gao, R. A. Shaikh and S. P. Soomro, proposed in " Recognition based segmentation of connected characters in text-based CAPTCHAs" an innovative approach to enhance recognition and segmentation of connected characters in text-based CAPTCHAs. To provide the necessary context, the authors conducted an exhaustive review of previous work related to recognition and segmentation of CAPTCHAs. The authors emphasized the significance of providing secure access to web applications, and how CAPTCHAs have evolved as a relevant security mechanism. Besides, they deliberated on the traditional machine learning algorithms used to achieve segmentation and recognition in CAPTCHAs, identifying the inadequacies therein. The authors proposed a novel recognition-based segmentation method for text-based CAPTCHAs that leverages morphological operations to reduce noise and separate connected characters. The experimental results demonstrated that their method provided superlative segmentation and recognition of text-based CAPTCHAs, leading to improved security in web applications[11].

Z. Fan discussed "Anti-cracking Technology for Captcha Recognition" , a new anti-cracking technology to enhance CAPTCHA recognition. To provide context, the author conducted a comprehensive literature survey on CAPTCHA recognition and anti-cracking technologies. The author discussed the importance of CAPTCHAs in providing web security and the limitations of traditional CAPTCHA systems. They also highlighted that attackers can use various methods to crack CAPTCHAs, compromising their security. The author proposed a new anti-cracking technology that works by generating dynamic CAPTCHAs that continuously update and change, mitigating the possibility of brute-force or dictionary attacks. The experimental results confirmed that the proposed approach is effective, robust, and reliable in combating CAPTCHA-cracking methods, making it a potential solution for web application security[12].

The paper by K. Qing and R. Zhang, named "A Multi-Label Neural Network Approach to Solving Connected CAPTCHAs" a novel approach to solving 'Connected CAPTCHAs using a multi-label neural network'. The authors conducted a literature survey on previous work related to solving CAPTCHAs using machine learning techniques. They discussed the limitations of traditional machine learning methods in solving CAPTCHAs, particularly for connected characters. They also deliberated on the significance of achieving accurate segmentation and recognition of connected characters in CAPTCHAs. To overcome these limitations, they proposed a multi-label neural network that combines a

convolutional neural network and a recurrent neural network. The experimental results demonstrate that the proposed method can accurately recognize connected characters in text-based CAPTCHAs, making it a promising solution for increasing the security of web applications [13].

Sr. no.	Authors	Year	Algorithm	Results	Conclusion
1.	Jing Wang, Jiaohua Qin *, Xuyu Xiang, Yun Tan and Nan Pan	2019	CNN , DenseNet-121 , ResNet-50	The DFCR model accuracy- 99%. Memory Consumption- 60%w.r.t. DenseNet having 80%. Training time- 3hrs.	1. They worked on 3 datasets including Chinese and English characters. 2. To increase the difficulty of recognition they designed images by rotating multiple characters. 3. They showed that the DFCR model is beneficial for feature reuse and compared the effectiveness with ResNet50 and DenseNet121[1].
2.	Zhong Wang and **Peibei Shi**	2021	CNN,AlexNet, VGG-16, GoogleNet, and ResNet	For dataset CNKI CAPTCHA, Zhengfang CAPTCHA, and randomly generated CAPTCHA, gave accuracy of 99%, 98.5% and 97.84% respectively	1. A new method for improving the traditional VGG network by using the CNN with focal loss function . 2. The focal loss function lessen the weight of samples that are simple to categorize and focus on hard-to-classify sample.[2]
3.	Y. Shu and Y. Xu	2019	CNN-multi, CNN-GRU, ResNet-multi, ResNet-GRU	Among mentioned algorithms, ResNet-GRU performs better resulting in 99.67%, 99.25% 99.875%,72.91% accuracy for Water Ripple,Shadow,FishEye and crawled dataset respectively.	1. They combine CNN with RNN to recognize CAPTCHA images without the need for prior segmentation and preprocessing. 2. Gated Recurrent Unit (GRU) is very similar to LSTM ,But it is much faster than LSTM in term of processing the same dataset.It outperform the LSTM when the dataset is small and has long text
4.	Zahra Noury , Mahdi Rezaei	2020	Deep Neural Network,	The proposed model achieved accuracy of 98.90% for Numerical data and 98.30% for Alphanumeric data	1. Adding certain characters like 3,8,g ,o,0 etc decreases the model accuracy while using brighter color characters also increases the difficulty level of the CAPTCHAs. 2. They compared the loss values of SGD and Adam optimiser for the training and testing process.
5.	Jun Chen, Xiangyang Luo, Yingying Liu,	2019	Deep Convolutional Neural Network	The proposed model achieved accuracy of detecting QQ captcha	1. They worked on only Chinesedataset. 2. To increase the identification

	Jinwei Wang and Yuanyuan Ma			is 95.4% and for detection Bot is 84.7%.	efficiency of confusion class and samples in confusion-class DCNN, a training and validating interactive learning technique is presented.
6.	Xiaohui Zhang, Xinhua Liu, Thompson Sarkodie-Gyan & Zhixiong	2019	Convolutional Neural Network (CNN)	The proposed model achieved accuracy of detecting Captcha on NetEase portal is 70.38% and for Baidu Tieba is 80.57% and for "Tuhu" and "QQ" is 81.24%.	1. For the deep learning model, a ten-layer network architecture was created, and three improved training methods were suggested. 2. A novel and quick technique was used to create a personalized Chinese character training set, with the goal of overcoming the limitations of labeled data collection and unequal data distribution.
7.	Jing Wang, Jiaohua Qin, Xuyu Xiang, Yun Tan and Nan Pan	2019	Deep Convolutional Neural Network	The recognition accuracy of CAPTCHA with the background noise and character adhesion is above 99.9%.	1. Improved and constructed a new DenseNet for CAPTCHA recognition. 2. Experiments demonstrate that the new network effectively lowers memory usage while maintaining the primary performance benefits of DenseNets.
8.	Y. Hu, L. Chen, and J. Cheng	2018	Deep Neural Network	Accurate recognition of CAPTCHAs with image and audio-based CAPTACHs.	1. The proposed approach can significantly enhance the accuracy of CAPTCHA recognition systems with deep learning. 2. Audio CAPTACHAs can also be cracked.
9.	G. An, W. Yu	2017	CAPTCHA Recognition Algorithm based on RSC-PPM	Accurate recognition of CAPTCHAs with accuracy with 95% of image and audio-based CAPTACHs.	1. Varying the technology of Deep Neural Networks can enhance accuracy significantly. 2. RSC-PPM algorithm is a promising solution to recognize connected CAPTCHAs and enhance security in web-based authentication.
10.	S. A. Kumar, N. R. Kumar, S. Prakash, K. Sangeetha	2017	Next Generation CAPTCHA with Gamification	Enhances internet security by increasing user engagement and providing rewards for participation.	1. Gamified CAPTCHAs have the potential to revolutionize internet security by improving user experience. 2. Technique used to provide increased security measures for web authentication.
11.	R. Hussain, H. Gao, R. A. Shaikh, S. P. Soomro	2016	Recognition depending Segmentation	Accurate segmentation of text-based CAPTCHAs. Average accuracy -	1. The proposed method of recognition-based segmentation provides precise segmentation of connected

				97%.	text-based CAPTCHAs. 2. Proposed method makes it a reliable solution for improving web security measures.
12.	R. Hussain, H. Gao, R. A. Shaikh, S. P. Soomro	2018	Recognition-based Segmentation using Morphological Operations	Provides accurate segmentation of connected characters in text-based CAPTCHAs, improving the security of web applications.	1. The proposed recognition-based segmentation method uses morphological operations to increase the precision of segmentation and recognition of text-based CAPTCHAs, 2. It makes it a promising solution for web application security.
13.	K. Qing, R. Zhang	2017	Multi-Label Neural Network Approach	Accurately recognizes connected characters in only text-based CAPTCHAs.	1. The proposed multi-label neural network approach can significantly improve connected CAPTCHA recognition accuracy and help enhance web security measures by defending against automated attacks. 2. Latency issues got fixed.

III. COMPARISON GRAPHS OF DIFFERENT ALGORITHMS

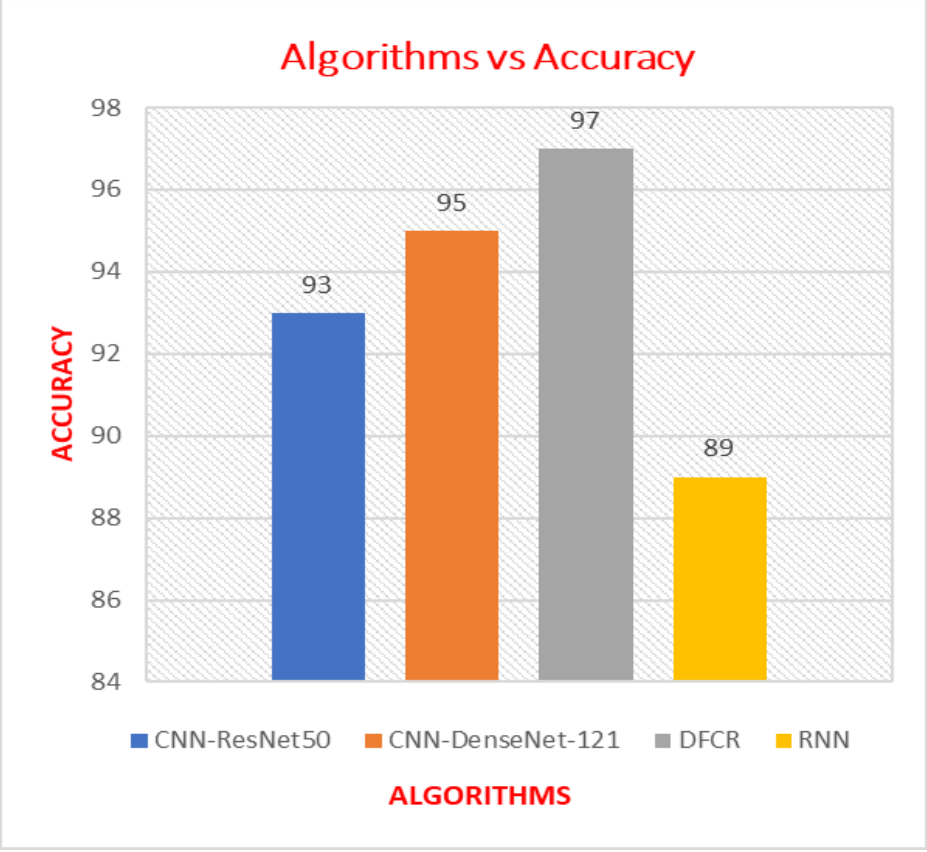

Fig.2 Accuracy Comparison of Existing Models

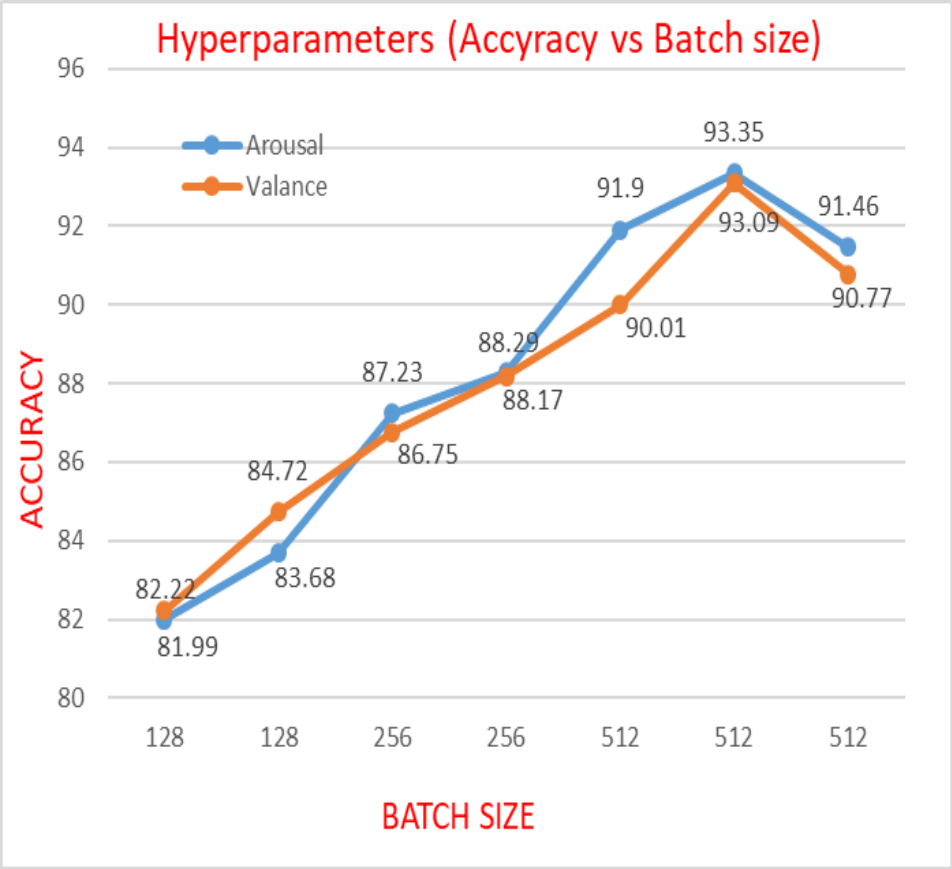

Fig.3 Batch Size vs Accuracy Comparison by Line Curve of (Hyperparameters Used)

IV. RESEARCH GAP

The requirement for models that can recognize captcha codes in various languages and formats is one of the key research gaps in deep learning-based captcha recommendation. Despite the fact that a sizable number of internet users use languages other than English to communicate, the majority of captcha recognition research concentrates on captcha codes in the English language. To ensure their efficacy across a wide range of web applications, captcha recommendation algorithms that can read and recognize captcha codes in various languages and formats should be developed. Models that can distinguish between real user behavior and automated bot activity are a critical research gap. Web security is compromised because sophisticated bots can read and access standard captcha codes. Hence, captcha recommendation models that can recognize humans and bots is now a need. Finally, research is required to assess the practical application of captcha recommendation models. While significant progress has been made in the technical aspects of captcha recognition using deep learning, the extent to which they improve web security and user experience remains inconclusive. These techniques are playing a very much important role in the field of digital cyber security. This concept lies in developing an accurate and robust model that can accurately recognize and classify complex Captchas with high accuracy, even when presented with noisy and distorted data Therefore, future research should address these gaps to provide more comprehensive solutions for captcha recommendation using deep learning.

V. CONCLUSION

The field of CAPTCHA recognition has seen significant advancements in recent years. The use of deep learning models for CAPTCHA recognition has shown great potential and has achieved high accuracy rates. This Literature review explores different methods to improve the recognition of CAPTCHA images, including the use of Dense Net, CNN with focal loss function, CNN with RNN, and selective learning confusion class for text-based CAPTCHA identification. Moreover, some researchers have studied ways to enhance the security of CAPTCHA, such as using specific numbers and digits that confuse the deep-CAPTCHA model with other letters and using brighter color characters. The proposed approaches show high accuracy rates and robustness in recognizing various styles of deformation of CAPTCHA images. Additionally, researchers have also proposed a novel approach based on deep learning to recognize audio-based CAPTCHA, making the technology more accessible to visually impaired users. Overall, deep learning models have proven to be effective in recognizing

CAPTCHA images and are an active area of research for future improvements.

VI. REFERENCES

[1] Wang, J., Qin, J., Xiang, X., Tan, Y. and Pan, N., 2019. CAPTCHA recognition based on deep convolutional neural networks. Math. Biosci. Eng, 16(5), pp.5851-5861.

[2] Wang, Z. and Shi, P., 2021. CAPTCHA recognition method based on CNN with focal loss. Complexity, 2021, pp.1-10.

[3] Y. Shu and Y. Xu, "End-to-End Captcha Recognition Using Deep CNN-RNN Network," 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2019, pp. 54-58, doi: 10.1109/IMCEC46724.2019.8983895.

[4] Noury, Z. and Rezaei, M., 2020. Deep-CAPTCHA: a deep learning based CAPTCHA solver for vulnerability assessment. arXiv preprint arXiv:2006.08296.

[5] Jun Chen, Xiangyang Luo, Yingying Liu, Jinwei Wang and Yuanyuan Ma.“Selective Learning Confusion Class for Text-Based CAPTCHA Recognition”2019, IEEE Xplore.

[6] Xiaohui Zhang, Xinhua Liu, Thompson Sarkodie-Gyan & Zhixiong. “Development of a character CAPTCHA recognition system for the visually impaired community using deep learning”, 2019, IEEE Xplore.

[7] Jing Wang, Jiaohua Qin, Xuyu Xiang, Yun Tan and Nan Pan. 2019 "CAPTCHA recognition based on deep convolutional neural network", 2019, Mathematical Biosciences and engineering.

[8] Y. Hu, L. Chen and J. Cheng, "A CAPTCHA recognition technology based on deep learning," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, 2018, pp. 617-620.

[9] G. An and W. Yu, "CAPTCHA Recognition Algorithm Based on the Relative Shape Context and Point Pattern Matching," 2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Changsha, China, 2017, pp. 168-172.

[10] S. A. Kumar, N. R. Kumar, S. Prakash and K. Sangeetha, "Gamification of internet security by next generation CAPTCHAs," 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2017, pp. 1-5.

[11] R. Hussain, H. Gao, R. A. Shaikh and S. P. Soomro, "Recognition based segmentation of connected characters in text-based CAPTCHAs," 2016 8th IEEE International Conference on Communication Software and Networks (ICCSN), Beijing, China, 2016, pp. 673-676.

[12] Z. Fan, "Anti-cracking Technology for Captcha Recognition," 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Huangshan, China, 2018, pp. 1032-1035.

[13] 6. K. Qing and R. Zhang, "A Multi-Label Neural Network Approach to Solving Connected CAPTCHAs," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 1313-1317.