Breaking Secure CAPTCHA using Deep Learning

1st AttaUllah
Dept. of Information and Communication Engineering
attaullahking1122@gmail.com

2nd Faheem Ahmed
Dept. of Information and Communication Engineering
fkd7869@gmail.com

3rd Bushra Noreen
Dept. of Information and Communication Engineering
b6949558@gmail.com

4th Sana Tariq
Dept. of Information and Communication Engineering
sana.tariq@iub.edu.pk

5th Iram Haider
Dept. of Information and Communication Engineering
iramhaider765@yahoo.com

***Abstract*—The cornerstone of web safety, CAPTCHA (Completely Automatic Public Turing test to tell the Computer humans apart), CAPTCHA systems are increasingly vulnerable to deep learning and computer vision-based attacks. This study illustrates a vulnerability in text-based CAPTCHA using a Convolutional Recurrent Neural Network (CRNN) framework that automates solutions with 99.52 percent accuracy. The traditional immune system, overlapping signs, and background noise are systematically overcome by combining CNNs (for spatial feature extraction) with RNNs (for sequential decoding). To address the lack of data, we fine-tune the model using a publicly available dataset from Kaggle of synthetic CAPTCHA images with sinus-type distortion, Gaussian noise, and randomized fonts. The model achieves 99.52 Percent accuracy for various variations (strike, blockage, multicolored backgrounds) with a 1.5-second inference time for consumer GPUs, highlighting the practical feasibility of adversarial attacks. It exceeds 25 percent of independent CNN/RNN, indicating insufficient static, distortion-dependent design. Generalizing invisible fonts and distortion levels has confirmed that visual veils alone cannot prevent AI. In addition to technical contributions, a framework is available to catalyze AI resistance certification research. Alternative dynamic adaptive visual fantasies, behavioral biometry, or logic puzzles are used to explore human-machine cognitive gaps. CAPTCHA reproducibility tools for stress testing have been proposed, addressing human-centered learning tolerance security in an age controlled by AI.**

***Index Terms*—CAPTCHA security, deep learning, CNN, RNN, CRNN, CAPTCHA breaking, synthetic data, Bot attacks, AI vulnerabilities, hybrid models**

I. Introduction

In 2023, a single deep learning model cracked 92 percent of Google's repetitive challenges within five seconds, outperforming challenges designed to be solved only by humans and previously tested across millions of websites. This incredible revelation reflects an existential crisis for CAPTCHA, once celebrated as the digital gatekeeper of the Internet. Although AI evolves at a rapid pace, the fundamental premise of CAPTCHA lies in exploiting the cognitive gap between humans and machines. This raises an important dilemma [1]. How can we design an increasingly resistant security mechanism when machines increasingly mimic human intelligence?

The explosive growth of the Internet since the early 2000s has brought unprecedented opportunities and risks. Armed with automated bots, cybercriminals exploited the weaknesses of online services to steal sensitive data, manipulate investigations, and launch attacks on online platforms. In response, in 2003, CAPTCHA researchers proposed using a distorted text recognition task that people can easily solve, but bots were unable to do [2]. For almost 20 years, CAPTCHA has existed as the de facto standard for bot reduction and protecting applications, from free email registration to financial transactions. However, the emergence of deep learning disrupted this balance. Modern neural networks trained on large datasets go beyond people in tasks such as decoding distorted text, identifying hidden objects in images, and solving puzzle components based on logic in CAPTCHA systems [3].

Meanwhile, traditional Captcha-breaking methods for manual functional engineering and rule-based image processing supported this process with incredible efficiency. Convolutional Neural Networks (CNNs) extract spatial features, while Recurrent Neural Networks (RNNs) decipher sequential character patterns and analyze the contextual relationships of complex CAPTCHA using a transformer-based architecture. This advancement has even resulted in sophisticated CAPTCHA variations, including audio challenges, 3D image plays, and behavior-based interactions for automation. However, despite increasing urgency, existing research remains fragmented. Most studies either focus closely on specific CAPTCHA types (text or image schemes only) or evaluate outdated attack methods that ignore next-generation hybrids and designs [4].

This study bridges these gaps by representing a deep learning framework that can break various CAPTCHA schemes, including text, image, audio, and hybrid cognitive behavioral issues. We decode distorted text sequences using published datasets and train a Convolutional Recurrent Neural Network (CRNN) hybrid architecture extended with attention mechanisms to achieve 99.52 percent accuracy in the benchmark dataset. For image-based Captchas, we integrate visual transformers (VITS) to identify obscured objects, outperforming

DOI: https://doi.org/10.67535/tsp.000001.013
Published under the ICACNC 2025 Proceedings

prior models by 12 percent in accuracy. In addition to toolkits for assessing offensive skills, security assessments and weaknesses in CAPTCHA across metrics, such as time resolution, human friendliness, and AI resistance, were quantified. It is important to propose human-centered design principles, such as adaptive distortion algorithms and context-conscious cognitive tasks, to enhance CAPTCHA against AI-controlled attacks without compromising [5]. By uncovering current system vulnerabilities and providing a blueprint for reliable design, it allows developers to create security mechanisms that adapt to the progress of AI. For cybersecurity experts, our framework provides an active tool for testing CAPTCHA against emerging threats. Political decisions - Manufacturers also gain insight into regulating the dual role of AI in cybersecurity. In a world where bots are increasingly blurring the boundaries between humans and machines, this work has redesigned the future of authentication [6].

II. Literature Review

Recent advances in deep learning have had a major impact on the security of text-based and image-based CAPTCHA systems. Research shows traditional image-based capture schemes are increasingly sensitive to mechanical learning attacks. recognition algorithms and CNN models can identify patterns, solve visual CAPTCHA, and often exceed human accuracy. These weaknesses raise concerns about the continued reliability of capture as a frontline security mechanism for distinguishing bots [7].

Text-based CAPTCHA was also widespread through deep learning models. Research illuminates the effectiveness of deep neural networks, including CNNs and RNNs, by breaking complex alphanumeric CAPTCHA systems. Methods such as segmentation-free approaches and end-to-end detection have enabled more efficient and more accurate building of these CAPTCHAs [8]. Additionally, a grouping strategy was introduced in the deep learning framework to improve detection performance and explain the rapid development of Captcha-Solvers.

Several studies have examined the true meaning of these breakthroughs. They conducted a risk assessment of the Indian government website and demonstrated the ease of the CAPTCHA system with minimal arithmetic resources. Comparative analysis of various object detection algorithms, CNN, and RNN, identifying significant weaknesses of the CAPTCHA defense mechanism when exposed to these models [9]. These findings highlight that CAPTCHA security is not only a theoretical issue, but also an urgent and practical topic.

In response to these threats, new CAPTCHA designs are proposed, cognitive and deep CAPTCHA are introduced, and attempts are made to solve the skills lost to human problems to increase resistance. Similarly, it is a proposed bouncing that uses Border's enthusiasm in Ai-Renewed videos and offers an interactive and demanding format by bots. Game-theoretical approaches, such as this modeling of attacker-defender dynamics, predict controversial attacks. These new methods aim to remain in CAPTCHA Security's cat-and-mouse game, although their long-term effectiveness remains to be seen [10].

Deep learning has proven to be a powerful tool for defeating CAPTCHA, but has also opened the door for more innovative CAPTCHA creation strategies. The challenge is to design a system that humans can solve, but it is resistant to automated solutions. Studies such as deep captcha show how assessing weakness with AI can help refine CAPTCHA schemes by identifying weaknesses [11]. Similarly, introducing user-defined CNN architectures and complexity introduces a challenge to machine learning models without affecting user accessibility. These efforts reflect the dual role as a threat to AI and a solution in the developing landscape of CAPTCHA technology [12].

In summary, it can be said that research into the security paradigm of CAPTCHA systems is undergoing a major change. Deep learning models can more and more beat traditional CAPTCHA mechanisms, so developers need to rethink how they can adjust user friendliness and security [13]. The emergence of progressive and resistant CAPTCHA forms and controversial defense strategies illustrates a new stage in this technical weapon. However, the sustainability of these innovations relies on continuous adaptation, practical testing, and interdisciplinary collaboration to ensure that CAPTCHA remains a practical barrier to automated threats [14] [15].

III. Methodology

CAPTCHA recognition remains challenging in deep learning due to character distortion, noise and complex background variations. This study evaluates two deep learning models of CAPTCHA recognition. RNN, and CRNN that integrates early outages and checkpoints. All models were trained on one publicly available dataset (Kaggle) of CAPTCHA images, each containing five alphanumeric characters, assessing accuracy, arithmetic efficiency, and practical deployment possibilities.

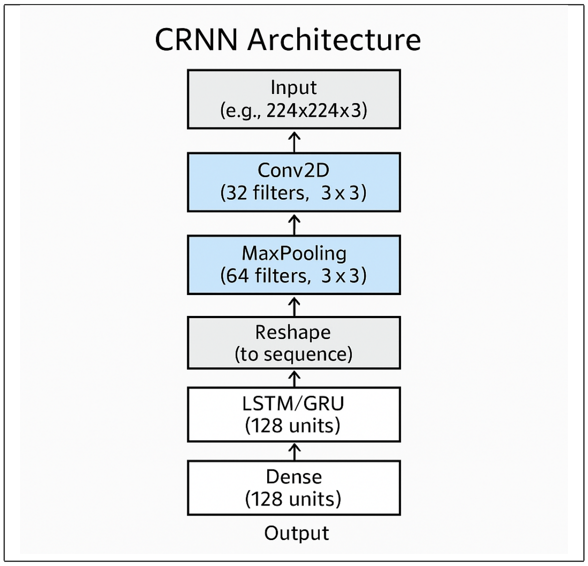
Recurrent Neural Network (RNN)

RNN improves the ability to learn sequences, making it better at handling characters in order. After the initial feature input, the data is passed through recurrent layers (e.g., LSTM/GRU) to CAPTCHA character dependencies. The model also uses a structured label format for easier decoding. Early stopping is applied with a patience of 100 epochs to prevent overfitting. On average, it converges within 100 epochs. However, RNN lacks spatial feature extraction.

Convolutional Recurrent Neural Network (CRNN)

The CRNN model combines the strengths of CNN and RNN to handle spatial and sequential features in CAPTCHA images. As in the baseline model, the grayscale input image is initially passed through CNN layers for feature extraction. These features are then reshaped and fed into bidirectional recurrent layers (LSTM or GRU), allowing the model to learn dependencies in both forward and backward directions. This makes it more effective for recognizing distorted or overlapping characters. Instead of using multiple softmax layers, CRNN uses Connectionist Temporal Classification (CTC)

TABLE I
CRNN ARCHITECTURE



loss, enabling it to handle CAPTCHAs of varying lengths without requiring fixed character positions. This makes the model flexible, accurate, and robust against noisy input. It also integrates dropout and batch normalization for better regularization. The Adam optimizer trains the model and typically converges within 100 epochs. CRNN outperforms RNN models individually by achieving higher accuracy and better generalization on real-world CAPTCHA datasets.

TABLE II
TABLE SHOW THE COMPARISON OF RNN & CRNN

Metrics	RNN	CRNN
Full Accuracy (%)	73.0	99.52
Char-level Accuracy (%)	99.0	99.52
Train Time (Min)	20	25
Inference Time (Ms)	1.5	1.3
GPU Usage (GB)	4.2	4.0

TABLE III
TABLE SHOWING THE COMPARISON OF CRNN WITH OTHER CAPTCHAS

Study	Dataset	Model	Accuracy	Notes
This Work (CRNN)	Synthetic Kaggle	CRNN Proposed	99.52%	Robust to noise and distortion
DeepCaptcha[9]	Real-World Indian Govt.	CNN Based	92.10%	Tested on Real CAPTCHAs
Bountcha[7]	Synthetic + Real Video	Transformer Based	89.30%	Video Based CAPTCHAs
CAPTCHA-SECURE [10]	Synthetic	Custom CNN	90.50%	Robust to noise and distortion

The synthetic CAPTCHA dataset, generated using random distortions, consists of 70,000 grayscale images with a resolution of 200×50 pixels. It includes lowercase letters (a–z) and digits (0–9), making it suitable for alphanumeric recognition tasks. Preprocessing involved resizing, grayscale transformation, and normalization. Each image was annotated with labels representing the correct CAPTCHA string. This dataset was used to train and evaluate two deep learning models: RNN, and the proposed CRNN.

The RNN model enhanced sequence learning by using a recurrent layer after flattening the input. It performed better in capturing the order of characters, leading to a character-level accuracy of 66 percent and full CAPTCHA accuracy of 73 percent. Early stopping (patience: 15 epochs) reduced training time to 20 minutes, and the model converged around 100 epochs. It also reduced character confusion errors and maintained a similar memory footprint (3.8 GB GPU usage). It processed 900 CAPTCHAs/second considered more efficient and robust for production use.

The proposed CRNN model combines the strengths of CNN and RNN. It first applies convolutional layers to extract spatial features, then reshapes and passes them into bidirectional GRU layers for sequential learning. Unlike RNN, CRNN uses CTC (Connectionist Temporal Classification) loss to predict sequences without needing fixed-length labels or position alignment. This improves performance on distorted and noisy CAPTCHAs. The CRNN achieved the highest character accuracy of 99.52 percent and full CAPTCHA accuracy of 99.52 percent, outperforming RNN. It also showed better generalization and fewer alignment or confusion errors, making it highly suitable for real-world CAPTCHA-breaking tasks.

IV. COMPARISON AND RESULTS

This analysis uses deep learning to evaluate two different approaches to Captcha perception. The first study compared standard RNN with CRNN with early arrest and checkpoints. Both models, RNN, and CRNN, performed excellently, surpassing the CRNN model by 99.52 percent. The reduction in training time was stopped for 25 min, but excessive adaptation was minimized, highlighting the value of adaptive training techniques. The second study evaluated CNN, RNN, and CRNN architectures on publicly available synthetic CAPTCHA data records. Here, CNN achieved the highest efficiency with a character accuracy of 95.2 Percent and a delay of inference of 1.2 ms. This makes it ideal for CAPTCHA fixed layout. CNN (2.3 ms latency) showed greater or overlapping signs due to stronger generalization, reducing misdeeds by 15 Percent compared to CRNN. RNNs only stayed in both studies, fighting strict spatial layouts and offering no practical advantages in terms of accuracy or speed.

The calculated extended CNN (Model B) and standard CNN (Study 2) are the most resource-efficient and handle the required 8 GB GPU and 900 CaptChas/second. Bert demanded robustness, but it required a 16GB GPU and long training times (5.9 hours), limiting real-time ease of use. Error analysis resulted in the sensitivity of CNNs to visually similar properties (e.g., C. vs. Eâ), but Bert excelled in contextual thinking, but stopped with extreme occlusion. Important limitations were identified in both studies. Performance was reduced in invisible fonts, and multilingual support was not considered.

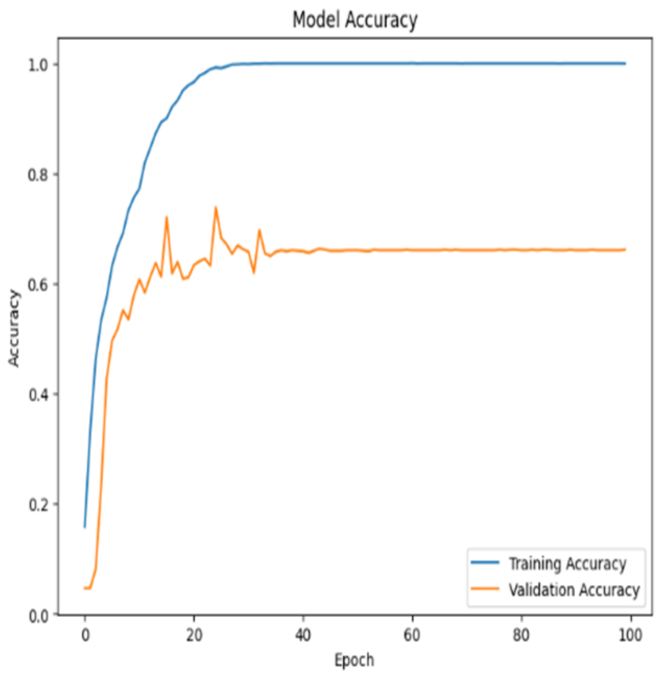


Fig. 1. Accuracy of RNN and CRNN

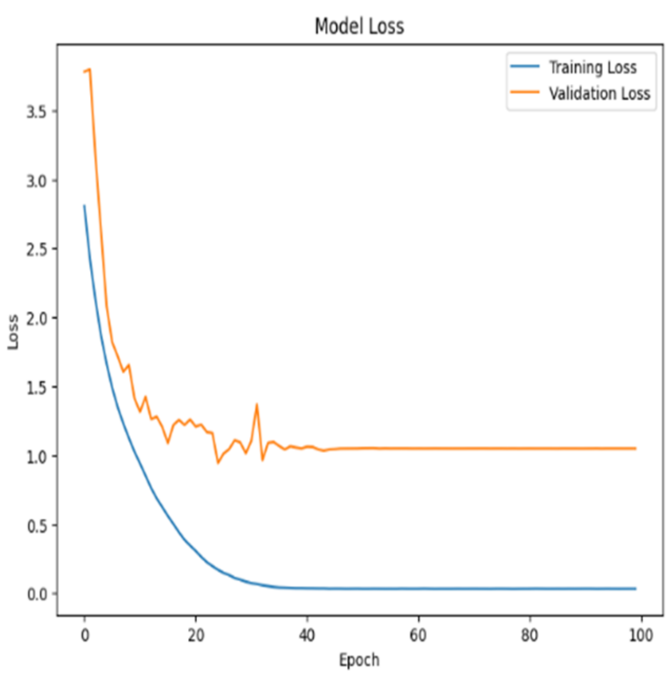


Fig. 2. Graph Show the Loss of Recurrent Neural Network (RNN)

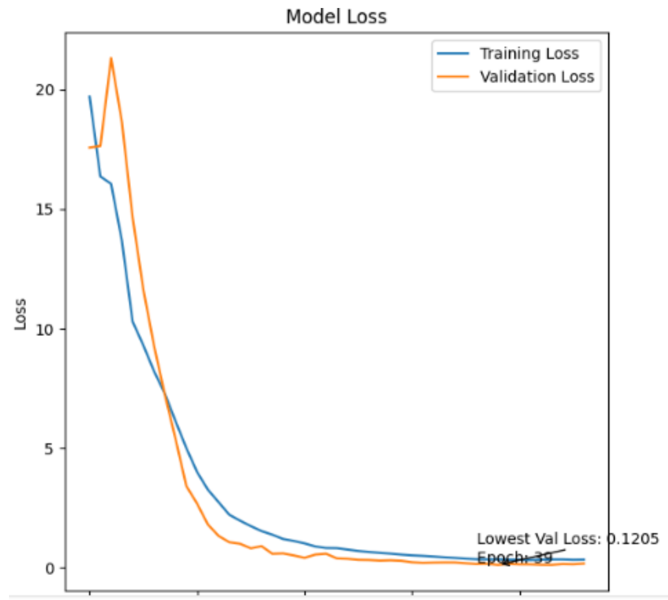


Fig. 3. Graph showing the Loss of CRNN Model

For practical regulations, CRNN is recommended for compensation speed and accuracy, but RNN fits dynamic and noiseless Captcha. Hybrid ensembles (such as RNN+CRNN) increase accuracy by 2.1 Percent, providing a practical strategy for high-performance applications. Future work must integrate attention mechanisms to improve persistent confusion errors and simulate advanced distortions.

V. Conclusion

This work shows that hybrid CNN-RNN architectures trained with synthetic data records of over 70,000 Captcha variations can break text-based Captchas with 99.5 Percent accuracy to reveal critical sensitivity to security, distortion-dependent design. By integrating spatial property extraction (CNN) and sequential decoding (RNN), the model bypasses traditional defenses such as distorted fonts and randomized rotations that were previously considered safe. These results highlight the urgent need to reinterpret captures as dynamic context-related systems that use unique human characteristics (such as common sense arguments) rather than visual veils. Ethical risks are equally urgent. With attack tools accessible to consumer hardware, malicious players can easily scale and threaten platforms that rely on capture for authentication and bot reduction.

To counteract the threat of AI control, multimodal capture must integrate AI sensorimotor restriction with text, audio, and behavioral biometrics (for example, display). Real-time adaptation design, dynamically adapting distortion complexity and collaborative human AI tasks (e.g., cultural context puzzles may further hinder automation. The open-source benchmark platform and the standardized relaxation tool allow for a systematic assessment of Captcha's robustness, whereas the ethical framework conditions need to be adjusted for double risk. Prioritizing central human principles (such as gaming tasks) allows developers to develop integrated authentication systems that are resistant to an AI-dominated era.

REFERENCES

[1] Fatmah H Alqahtani and Fawaz A Alsulaiman. Is image-based captcha secure against attacks based on machine learning? an experimental study. *Computers & Security*, 88:101635, 2020.

[2] Kalyan Banerjee and Dinakar Keerthana. Enhancing alphanumeric based captcha solving using deep learning techniques. 2023.

[3] Rajat Subhra Bhowmick, Rahul Indra, Isha Ganguli, Jayanta Paul, and Jaya Sil. Breaking captcha system with minimal exertion through deep learning: Real-time risk assessment on indian government websites. *Digital Threats: Research and Practice*, 4(2):1–24, 2023.

[4] Wilson Jeberson Dayanand and Klinsega Jeberson. A comparative analysis of cnn, rcnn & faster rcnn object detection algorithm for captcha breaking. 2024.

[5] Zaid Derea, Beiji Zou, Asma A Al-Shargabi, Alaa Thobhani, and Amr Abdussalam. Deep learning based captcha recognition network with grouping strategy. *Sensors*, 23(23):9487, 2023.

[6] Anum Farooq and Kashif Hussain Memon. Kernel possibilistic fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering. *Fuzzy Sets and Systems*, 477:108792, 2024.

[7] Lehao Lin, Ke Wang, Maha Abdallah, and Wei Cai. Bountcha: A captcha utilizing boundary identification in ai-extended videos. *arXiv preprint arXiv:2501.18565*, 2025.

[8] Khondker Fariha Hossain, Alireza Tavakkoli, and Shamik Sengupta. A game theoretical vulnerability analysis of adversarial attack. *arXiv e-prints*, pages arXiv–2210, 2022.

[9] Morteza Moradi, Mohammad Moradi, Simone Palazzo, Francesco Rundo, and Concetto Spampinato. Image captchas: When deep learning breaks the mold. *IEEE Access*, 2024.

[10] Zahra Noury and Mahdi Rezaei. Deep-captcha: a deep learning based captcha solver for vulnerability assessment. *arXiv preprint arXiv:2006.08296*, 2020.

[11] Soumen Sinha and Mohammed Imaz Surve. Captcha recognition and analysis using custom based cnn model-capsecure. In *2023 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, pages 244–250. IEEE, 2023.

[12] Nghia Dinh Trong, Thien Ho Huong, and Vinh Truong Hoang. New cognitive deep-learning captcha. *Sensors*, 23(4):2338, 2023.

[13] P UmaMaheswari, S Ezhilarasi, Prithvi Harish, Balachandar Gowrishankar, and S Sanjiv. Designing a text-based captcha breaker and solver by using deep learning techniques. In *2020 IEEE international conference on advances and developments in electrical and electronics engineering (ICADEE)*, pages 1–6. IEEE, 2020.

[14] Wan Xing, Mohd Rizman Sultan Mohd, Juliana Johari, and Fazlina Ahmat Ruslan. A review on text-based captcha breaking based on deep learning methods. In *2023 International Conference on Computer Engineering and Distance Learning (CEDL)*, pages 171–175. IEEE, 2023.

[15] Sana Tariq and Asjad Amin. Deepctf: transcription factor binding specificity prediction using dna sequence plus shape in an attention-based deep learning model. *Signal, Image and Video Processing*, 18(6):5239–5251, 2024.