Project Proposal
AI-powered CAPTCHA Solver

Rohan Javed (L21-5625) and Muhammad Mahd Kashif (L20-1299)
FAST National University of Lahore (FAST NU LHR)

March 2024

Introduction and Motivation

The prevalence of CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) has become a ubiquitous hurdle in online interactions. These challenges aim to distinguish automated bots from human users. While essential for security purposes, CAPTCHAs can be frustrating and time-consuming to solve, especially for users with visual impairments or slow internet connections.

This project proposes an AI-powered CAPTCHA solver that leverages machine learning techniques to automate the process. This approach holds significant potential for:

- Enhanced User Experience: Users can bypass the CAPTCHA challenge efficiently, saving time and frustration.

- Improved Accessibility: Users with visual impairments can rely on the AI to interpret and solve CAPTCHAs.

- Security Automation: Security systems can integrate the solver to automate CAPTCHA challenges encountered during automated tasks.

Challenges

Developing an effective AI-powered CAPTCHA solver presents several technical hurdles:

- CAPTCHA Diversity: CAPTCHAs come in various forms, including text-based, image-based, audio-based, and even interactive challenges. The solver needs to be adaptable to handle diverse CAPTCHA types.

- Evolving CAPTCHAs: CAPTCHA design constantly evolves to counter automated solvers. The system must be able to learn and adapt to new CAPTCHA variations.

- Balancing Accuracy and Speed: The solver requires high accuracy to bypass CAPTCHAs successfully, while maintaining efficiency to minimize delays.

State of the Art and Differentiation

Existing CAPTCHA-solving solutions often rely on rule-based approaches or simple image recognition techniques. These methods struggle with complex CAPTCHAs and lack the ability to adapt to evolving designs. Recent research explores applying deep learning techniques for CAPTCHA solving. However, these approaches can be computationally expensive and require large training datasets.

Our project differentiates itself by:

- Hybrid Approach: Combining image recognition, text recognition, and machine learning techniques to handle diverse CAPTCHA types.

- Continuous Learning: Implementing a self-learning mechanism to adapt to new CAPTCHA variations through continuous training on encountered CAPTCHAs.

- Focus on Efficiency: Optimizing the AI model for fast and lightweight operation, minimizing latency during CAPTCHA solving.

High-Level Architecture

The proposed CAPTCHA solver will consist of the following modules:

- CAPTCHA Preprocessing: The received CAPTCHA will be preprocessed for noise reduction and format conversion.

- CAPTCHA Classification: A machine learning model will classify the CAPTCHA type (text-based, image-based, etc.).

- Solution Module: Based on the classification, specific modules will be employed to solve the CAPTCHA (e.g., Optical Character Recognition (OCR) for text-based CAPTCHAs, image segmentation and recognition for image-based CAPTCHAs).

- Validation: The solution will be validated against a human-in-the-loop mechanism or a confidence score to ensure accuracy.

Datasets

The project will utilize publicly available CAPTCHA datasets, potentially including:

- OpenAI CAPTCHA Dataset: `https://help.openai.com/en/articles/8184038-captchas-in-chatgpt`

- HiCaptcha Dataset: `https://github.com/topics/hcaptcha?o=desc&s=updated`

These datasets provide a variety of CAPTCHA types for training and testing the AI model.

References

Brueggemann, S., Liu, Y. (2016). CAPTCHAs: Usability challenges and breaking the paradox. IEEE Transactions on Human-Machine Systems, 46(5), 632-645. `https://ieeexplore.ieee.org/document/9544570`

Conclusion

This proposal outlines a project that leverages AI techniques to develop a robust and adaptable CAPTCHA solver. By addressing the technical challenges and utilizing a hybrid approach, this project aims to improve user experience, accessibility, and automation capabilities in the digital world.