ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Breaking Image CAPTCHAs and designing a new CAPTCHA

Question of the thesis. What causes adversarial examples and how can they be used to design CAPTCHAs?

CAPTCHA In case you have ever tried parsing websites for data or just surfed the internet, the odds of you encountering some type of CAPTCHA system are higher than odds of you understanding the text, that is written here. The work on bypassing CAPTCHAs should be seen as a motivation to develop new better ways to ensure proper bot isolation from the resources and materials designed for human use. There are CAPTCHA solving service-providers claiming to be able to do this reliably, so potentially one can benchmark and compare solutions in this area, including the one that will be born in this research. Based on experience and data of previous projects, develop tool uniting CAPTCHA solvers from different providers. The goal output of this study is a deployable plugin/app that allows solving CAPTCHAs as well as a new method of testing users for "humanness" as well as potentially more AI resistant protection system.

Modelling Guidelines It is advised to start exploration with lighweight models (e.g. YOLO), but the only requirement is that models applied in the project must have open architecture (open source). Mechanisms differentiating CAPTCHA providers and task types do not need to be present in the beginning. For the generation of new CAPTCHA techniques same suggestions apply.

1 Data Collection and Annotation

1.1 CAPTCHA solver

In the first stage of this project, It is necessary to test performance of pre-trained readily available model instances on open source platforms. Upon this search for or collection of a dataset to improve accuracy is expected.

Time Estimate Data collection and preparation per CAPTCHA source should not exceed 1 week (including known task types).

1.2 CAPTCHA generator

In the second stage of this project, It is necessary to test performance of people with diverse background and experience on the generated tasks. This part is to be in sync with the Data Collection section.

Time Estimate Data collection should not exceed 1 week with a balanced age based grouping.

2 Choosing a Suitable Metric for Automatic Evaluation

Evaluation Metric Success rate for a solution is the standard for the evaluation. For the first stage several flagged settings including enabling user's mouse movement as well as single key strokes are to be considered separately. Other parameters and data representation are up to the researcher upon discovery of a valuable insight reflection. For the second stage the time required versus success rate are to be compared for age based groups of people.

Validation The solution can be validated on the in-house test environment that will need to be developed as well.

3 Experimental Baselines

Success rate is to be compared with results acquired from human interactions. This part will be assisted with throughout the study.

4 Further Directions

CAPTCHA dev trends The main purpose of the second stage is to compare possible directions of CAPTCHA development to improve human-bot classification. Related findings are to be reported, considering the assumptions and limitations of the conducted study, summarizing possible improvements in the Conclusion section.

Detailed Project Outline

We denote the following primary tasks mandatory (on the right side you find a rough estimate for the time that we allocate to the respective task), however the direction of the project is flexible:

- Literature review (related work, previous approaches) ($\star$)
- Understanding required techniques and libraries (e.g. 3D image processing, graph construction frameworks) ($\star$)
- Assessing previous image analysis blood vessel graph reconstruction pipelines ($\star\star$)
- Further research and implementation of ML approaches ($\star\star\star\star$)
- Verification and validation of the graph representation ($\star\star$)
- Evaluating performance against existing baselines ($\star$)
- Writing the final report ($\star\star\star$)
- Midterm and final presentations ($\star$)

Extensions

- Investigate and implement more sophisticated methods for determining edge weights and/or other graph properties (e.g. directionality)
- Perform a detailed comparative study between the developed methods and existing techniques for blood vessel network analysis
- Verify the pipeline by running network simulations

The Student's Duties

- One meeting per week with the advisers to discuss current matters
- A final report in English, presenting work and results
- A midterm and a final presentation (15 min) of the work and results obtained in the project