Breaking CAPTCHA

Context

- CAPTCHA attempts to provide security against bots and can appear in many forms, including text, image, audio and video.
- Initially, it was realized that CAPTCHAs could be used for security purposes because there were no AI programs that could resolve the problem task.
- We want to evaluate the current state of the art of AI methods to break CAPTCHA in order to determine the validity of this assumption.

Project

- The core idea of the project is to use deep-learning methods and train a model that breaks image-based CAPTCHAs
- The project should start with an overview of the different types of CAPTCHAs and of the protection methods (distortion, rotation, noise, etc.)
- The students should choose a target CAPTCHA, generate two databases (training and testing) and train a model that will break the target. In order to achieve this, the students should read the proposed scientific papers, choose a method and apply it.

Scientific papers

- CAPTCHA recognition with Active Deep Learning, German Conference on Pattern Recognition Workshop 2015
- Yet another text CAPTCHA solver, CCS 2018
- Text-based CAPTCHA Strengths and Weaknesses, CCS 2011
- A simple generic attack on text captchas. NDSS 2017
- A Low-cost Attack on a Microsoft CAPTCHA, CCS 2008
- CAPTCHA Recognition Using Deep Learning with Attached Binary Images, Electronics 9, 2020