Solving Image-Based CAPTCHAs Using Optical Character Recognition and Machine Learning

Fogel, Matthew (School: Yucaipa High School)

The purpose of this project was to test the effectiveness of image-based CAPTCHAs against Optical Character Recognition (OCR) and machine learning algorithms. The design constraints were for the program to require no human intervention aside from starting the tasks and for the program to solve CAPTCHAs from start to finish before the CAPTCHA times out (approximately 60 seconds). First, a Graphical User Interface (GUI) containing controls and a custom console widget was designed and programmed. Next, browser automation capabilities were added with the Python Selenium library in combination with the Firefox Gecko webdriver. This allowed the program to perform actions within the web browser. The PyTesseract library was added to perform OCR on a screenshot of the prompt. The OCR was run at three separate zooms to retrieve the most accurate results. Finally, TensorFlow was added to the program to detect objects within the CAPTCHA image. After multiple tests to determine the fastest but most accurate model type, the program was modified to determine the coordinates of the objects within the images. The program was then completed to determine which boxes on the grid the coordinates belonged to and, finally, click those boxes. In the end, the program could solve almost every type of image-based CAPTCHA, with the exception of a few. On average, the three zooms of OCR took 10.945 seconds, and the object detection took 11.582 seconds. With the tasks that it could solve, the program met the design constraints.

Awards Won: