

Readme

Since the test cases are texts in the form of an image, these captchas could be bypassed using the Optical Character Recognition (OCR) technology that recognizes the text inside images, such as scanned documents and photographs.

I used tesseract which is an open-source text recognition (OCR) Engine and used the library pytesseract to retrieve the text in the image that is loaded from the URL which is scraped from the webpage using selenium and then used functions of selenium to find the form elements and sent the solution text from tesseract module to the form box. At the last, click was simulated at the submit button to submit the form.

Note: The chrome browser shuts down after the execution of code so `time.sleep(30)` makes the code wait for 30 seconds before cleaning up memory so that the results can be observed properly.