

Python 데이터 분석과 이미지 처리

나동빈

중급 Captcha Hacking 2 – 데이터 수집과 분석

수집할 데이터의 개수

- Q. 데이터를 많이 수집해야 할까요?
- A. 이미지 내의 문자 형태가 거의 동일하므로, 데이터를 조금만 수집해도 될 것 같습니다.

중급 Captcha Hacking 2 – 데이터 수집과 분석

데이터 수집하기

- 데이터의 개수가 적으므로 브라우저로 직접 수집합니다.
- 가능한 12개의 문자: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -, *
- 대략적으로 최소한 5개 이상의 이미지가 필요합니다.



중급 Captcha Hacking 2 - 데이터 수집과 분석

데이터 분석하기

- 색상 추출기를 활용하여, 각 문자의 색상이 어떻게 구성되어 있는지 확인합니다.



중급 Captcha Hacking 2 – 데이터 수집과 분석

데이터 분석하기

- 파란색: RGB에서 B 값이 항상 FF입니다.
- 초록색: RGB에서 G 값이 항상 FF입니다.
- 빨간색: RGB에서 R 값이 항상 FF입니다.
- 파란색 & 초록색: RGB에서 R 값이 항상 AA 이하입니다.
- 파란색 & 빨간색: RGB에서 G 값이 항상 AA 이하입니다.
- 초록색 & 빨간색: RGB에서 B 값이 항상 AA 이하입니다.