

Quebrando CAPTCHAs

Julio Trecenti

17 de maio de 2018

Sumário

1	Introduction	4
2	Introdução	5
2.1	Objetivos	5
2.2	Resultados Esperados	5
2.3	Organização do trabalho	5
3	Problema	6
3.1	Variantes	7
3.1.1	Áudio	7
3.1.2	Covariáveis e número de respostas variável	7
3.1.3	reCaptcha	7
4	Solução	8
5	Resultados	9
6	Considerações finais	10
7	Pacote decryptr	11
8	CAPTCHAs em áudio	12

Lista de Tabelas

Draft

Lista de Figuras

Draft

Capítulo 1

Introduction

Draft

Capítulo 2

Introdução

2.1 Objetivos

2.2 Resultados Esperados

2.3 Organização do trabalho

Capítulo 3

Problema

O problema do Captcha pode ser entendido como um problema de classificação de imagens. Especificamente, nosso interesse é criar uma função g que recebe uma imagem $\mathbf{X} = \{x_{nmr} \in [0, 1]\}_{N \times M \times R}$ e retorna um vetor de índices \mathbf{y} , sendo que cada índice y_j corresponde a um caractere c_j , $j = 1, \dots, L$, onde L é o número de caracteres contidos na imagem.

Das afirmações anteriores podemos tirar três conclusões.

1. Nossa variável **explicativa**, a imagem, é uma matriz $\mathbf{X} = \{x_{ijk}\}_{N \times M \times R}$, em que N é o número de linhas, M é o número de colunas e R é o número de *cores*, ou *canais*.

O elemento x_{nm} é denominado *pixel*. Um pixel representa a menor unidade possível da imagem. Em uma imagem colorida, por exemplo, temos $R = 3$. Nesse caso, um pixel é um vetor de três dimensões com valores entre zero e um, representando a intensidade de vermelho, verde e azul da coordenada n, m da imagem. Numa imagem em escala de cinza, temos $R = 1$ e o pixel, de uma dimensão, representa a intensidade do cinza (com 1=branco e 0=preto).

2. O objeto $C \in \mathcal{A}^L$ é um vetor de itens de um alfabeto \mathcal{A} com tamanho $|\mathcal{A}|$, finito e conhecido. Esse alfabeto contém todos os possíveis caracteres que podem aparecer na imagem.
3. Nossa **resposta** $\mathbf{y} \in \{1, \dots, |\mathcal{A}|\}^L$ é um vetor de índices de tamanho fixo. Cada elemento de \mathbf{y} representa um valor do alfabeto \mathcal{A} .

A construção de uma função g capaz de mapear \mathbf{y} a partir de uma nova imagem \mathbf{X} depende de uma amostra de imagens $\mathbf{X}_1, \dots, \mathbf{X}_S$ corretamente classificadas por $\mathbf{y}_1, \dots, \mathbf{y}_S$. A tarefa é, portanto, estimar uma função \hat{g} com o objetivo de minimizar.

$$L(g(\mathbf{X}), \mathbf{y}) = \mathbb{I}(g(\mathbf{X}) \neq \mathbf{y})$$

em que \mathbb{I} é a função indicadora.

3.1 Variantes

3.1.1 Áudio

Captchas também podem se manifestar na forma de áudio. Nesses casos, o usuário é condicionado a ouvir um áudio e transcrever seu conteúdo em um texto.

Com base nos Captchas analisados durante o desenvolvimento do trabalho, verificamos que Captchas de áudio são menos complexos. Por exemplo, alguns destes são formados por sons sem ruído. Ou seja, uma tabela de sons e classificações seria suficiente para quebrar os Captchas.

Nos casos com ruído, podemos utilizar duas técnicas para quebrar Captchas de áudio. A primeira é baseada em engenharia de características [ref], que extrai covariáveis dos áudios, para serem utilizados posteriormente em um modelo de regressão clássico. O segundo método consiste em calcular o espectrograma do áudio e tratá-lo como um Captcha de imagem.

3.1.2 Covariáveis e número de respostas variável

3.1.3 reCaptcha

Capítulo 4

Solução

Draft

Capítulo 5

Resultados

Draft

Capítulo 6

Considerações finais

Draft

Capítulo 7

Pacote decryptr

Draft

Capítulo 8

CAPTCHAs em áudio

Draft