

# Quebrando CAPTCHAs

Julio Trecenti

03 de maio de 2018

# Sumário

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

## 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**

### **3.1.2 Covariáveis e número de respostas variável**

### **3.1.3 reCaptcha**

Draft



## Capítulo 4

### Solução

Draft

## **Capítulo 5**

### **Resultados**

Draft

## **Capítulo 6**

### **Considerações finais**

Draft

## Capítulo 7

### Pacote decryptr

Draft

## Capítulo 8

### CAPTCHAs em áudio

Draft