



Disciplina Introdução à Computação Visual	Turmas
Professor William Robson Schwartz	

Entrega: 14/01/2025 até às 23h59 (via moodle), apresentações dias 15/01 e 20/01

## Trabalho Prático 2: Reconhecimento de Captcha

O objetivo deste trabalho é reconhecer captchas como os mostrados na Figura 1. *O trabalho deve ser feito individualmente ou em duplas.*



Figura 1: Exemplos de imagens de entrada.

### O que deve ser feito

Implementar **dois métodos** que, dada uma imagem de entrada como aquelas mostradas na Figura 1, reconheçam cada um dos seis caracteres do captcha. O primeiro método deve considerar o descritor HOG visto em sala de aula e um classificador de sua escolha. O segundo método deve considerar o uso de redes convolucionais. Embora seja esperado que o segundo método obtenha melhores resultados, escolhas cuidadosas feitas para o primeiro método (HOG + classificador), podem levar à obtenção de resultados apenas um pouco abaixo daqueles obtidos pelo segundo método.

As imagens contendo os captchas são disponibilizadas em três diretórios: *treinamento*, *validacao* e *teste*. Qualquer aprendizado deve ser feito apenas com a utilização das imagens contidas no diretório *treinamento* e qualquer validação de parâmetros deve ser feita apenas com imagens do diretório *validação*. As imagens do diretório *teste* não podem ser utilizadas nem para o aprendizado nem para validação, devem ser apenas utilizadas para avaliação do método (descrita abaixo).

Após o treinamento e validação, o seu método deve carregar cada uma das imagens do diretório *teste* e executar o reconhecimento de cada um dos caracteres da imagem e gerar uma saída que mostra a imagem acima de uma string com os caracteres reconhecidos, como exemplificado na Figura 2.

Ao final, também gere um gráfico mostrando a taxa de reconhecimento em função do número mínimo de caracteres que foram corretamente reconhecidos por captcha, de acordo com o reconhecimento correto (ground-truth) disponível no diretório labels10k. Um exemplo deste gráfico é mostrado na Figura 3. O exemplo da Figura 3 indica que apenas 40% das vezes três ou mais caracteres de um captcha são reconhecidos corretamente e que apenas 10% das vezes todos os seis caracteres de um captcha são reconhecidos corretamente. O gráfico ideal seria aquele onde todos os valores do eixo x correspondem ao valor 1 no eixo y (todos os caracteres estão sendo reconhecidos corretamente).

### O que deve ser entregue

Um relatório com no **máximo três páginas** contendo explicação sobre os métodos implementados e análises dos resultados obtidos com ambos os métodos (com gráficos, como o da Figura 3, e resultados



Figura 2: Exemplo da saída a ser gerada (mostra cada imagem do conjunto de teste e a uma string de com os caracteres reconhecidos). *Note que a string a ser mostrada é a string reconhecida pelo método proposto e não a string correta (ground-truth), que está disponível no diretório labels10k.*

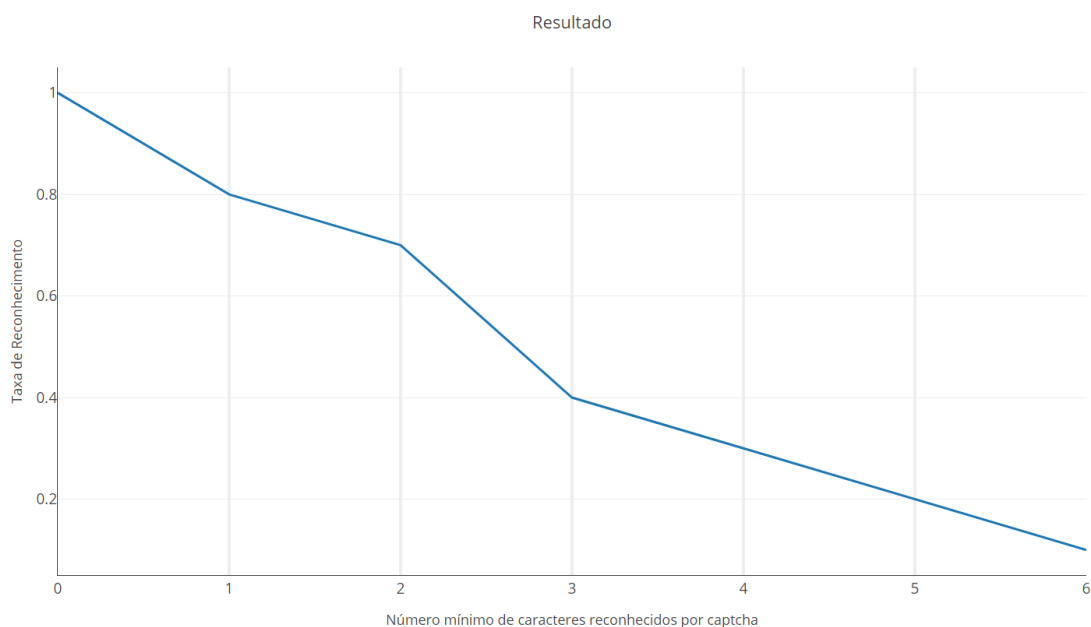


Figura 3: Exemplo do gráfico a ser gerado.

qualitativos, como mostrado na Figura 2), assim como uma discussão sobre prós e contras de cada método implementado.

## O que será disponibilizado

- Conjunto de imagens: imagem contendo as imagens nas seus respectivos diretórios e um diretório com o *ground-truth* contendo os caracteres corretos para cada captcha, de acordo com o nome do arquivo.

## **Critérios para Correção**

- Engenhosidade: 20% - Criatividade no desenvolvimento da solução (deve estar documentado no relatório e na apresentação).
- Resultados: 40% - Resultados e análises dos resultados obtidos (deve estar documentado no relatório e na apresentação).
- Relatório 20% - Aspectos incluindo contextualização do problema, solução proposta, resultados, discussão, estrutura, clareza e corretude.
- Apresentação: 20% - Aspectos incluindo estrutura, contextualização do problema, solução proposta, resultados obtidos, análise comparativa dos resultados e fluxo da explicação.