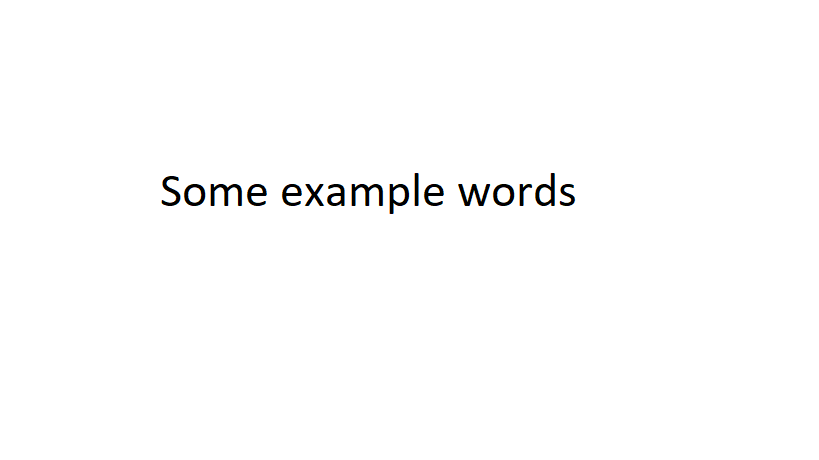
**Projeto 2**

**Exercício 1**

1. Search online by an Optical Character Recognition (OCR) software tool, https://en.wikipedia.org/wiki/Optical\_ character\_recognition, developed on any programming language at your choice. Test the software and check its proper functioning, on some images at your choice. Report on the experimental results.

Foi utilizado o método ocr(…) do MATLAB. Este método recebe uma imagem e devolve um objeto que contém o texto que encontrou na imagem, para além de outras informações.

Utilizando a seguinte imagem:

O output obtido foi:

“Some example words”

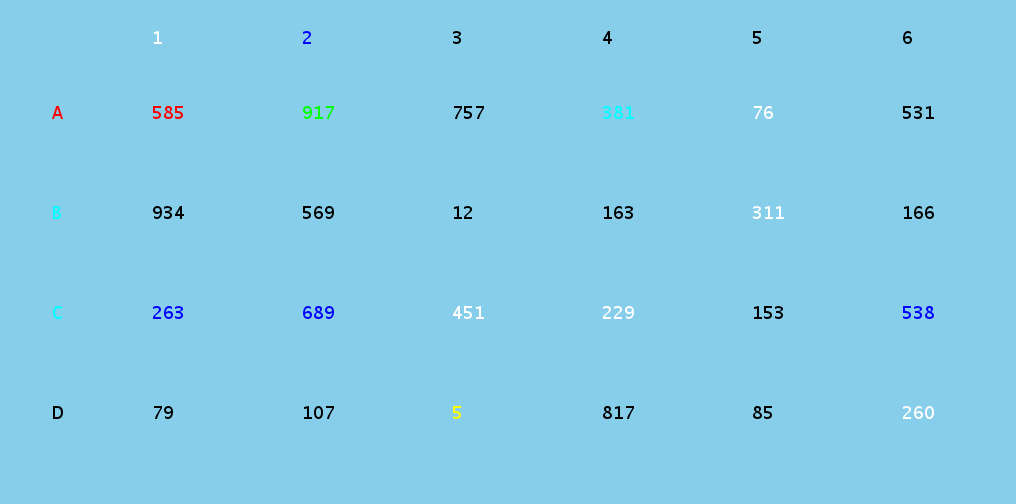
1. Consider the grayscale and color images generated with the CodeCardGenerator and CAPTCHA applications. Apply the chosen OCR software to these images. Comment on the results regarding the obtained text extraction quality. Identify in which situations and conditions in the input image, the OCR software:
   * extracts the text with no errors;
   * extracts the text with some errors;
   * fails to extract (e.g. ignores) the existing text;

**Extração sem erros:**



O algoritmo consegue distinguir facilmente os números e letras do fundo e por isso consegue obter as palavras na sua totalidade.

**Extração com alguns erros:**

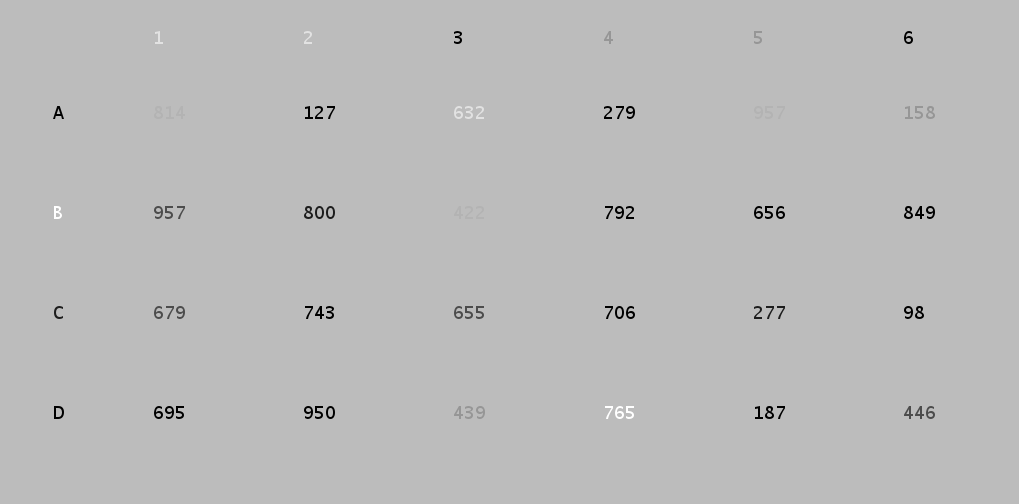
Texto extraído:

535, 934, 263, 79,

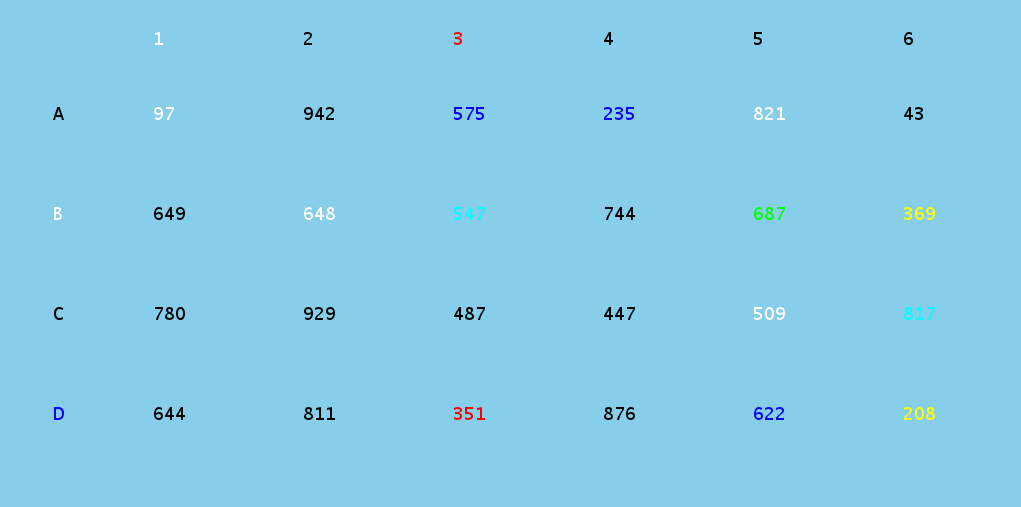
569, 689, 107, 757,

12, 163, 817, 153,

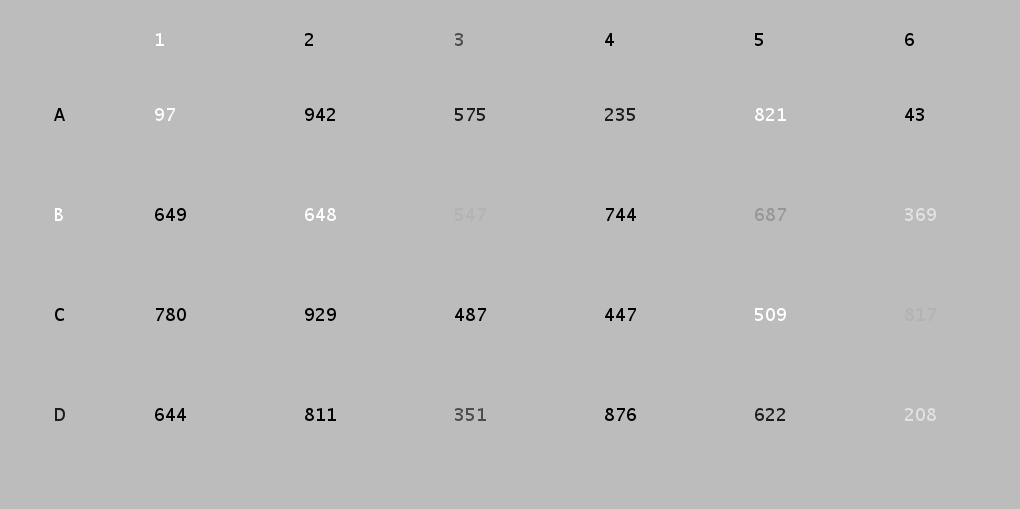
85, 531, 166, 538

Texto extraído:

957, 679, 695, 127, 800, 743, 950, 655, 279, 792, 706, 656, 277, 187, 849, 98, 446

Texto extraído:

649, 780, 644, 942, 929, 811, 575, 487, 351, 235, 744, 447, 876, 622, 43

Texto extraído:

649, 780, 644, 942, 929, 811, 575, 487, 351, 235, 744, 447, 876, 622, 43

Texto extraído:

IHXLIAB

Analisando as imagens e os respetivos resultados, percebemos que o algoritmo consegue identificar alguns dos números nas primeiras imagens, mas não todos.

Os resultados são significativamente melhores nos casos onde o contraste entre o número e o fundo é maior. Uma vez que os fundos são claros, o desempenho é melhor quando os números são mais escuros.

Na última imagem, que está desfocada, o algoritmo conseguiu obter as primeiras letras da palavra, mas falhou em obter as últimas corretamente.

**Falha na extração:**



Devido ao facto das letras e números da imagem estarem desfocados, o algoritmo foi incapaz de obter um único número ou letra.