

### Lista de Exercícios No. 1

1. Deseja-se projetar um sistema para leitura automática das placas de veículos em um estacionamento. Produza um roteiro para a solução do problema, caracterizando as etapas do processo de visão computacional, da aquisição ao reconhecimento. Em cada etapa, descreva uma técnica que poderia ser utilizada para a solução deste problema específico.
2. Diferencie amostragem e quantização
3. Na série de TV “The designated survivor” (Netflix) uma detetive tenta recuperar a impressão digital em uma taça de vinho a partir de uma foto obtida com a câmera de um smartphone. Sabe-se que os picos e vales de uma impressão digital têm em média 0.2mm de espessura. Considerando que a foto capturou um campo visual de 1 m<sup>2</sup>, qual teria que ser a resolução mínima da câmera para isso ser possível?
4. Porque o problema da segmentação é considerado sem solução geral?
5. Para cada imagem abaixo:

A

3	5	2	1	1
1	4	6	2	1
1	1	5	6	2
1	1	1	1	1
1	2	2	2	1

B

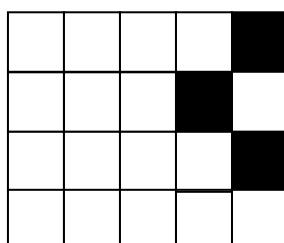
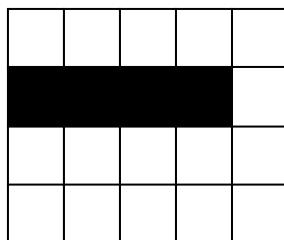
5	1	2	1	8
6	6	5	6	1
2	1	8	7	7
6	1	2	8	8
7	8	2	1	1

C

1	1	9	1	1
1	1	9	8	7
9	9	9	2	1
1	1	2	8	8
1	2	2	8	9

- a) Determine sua resolução e quantização.
- b) Segmente a imagem pela técnica de limiar, com valor de corte  $t > 2$  para A,  $t > 5$  para B e  $t > 6$  para C.
- c) Aplique o algoritmo de rotulação no resultado do item anterior. Diga como fica a tabela de rótulos equivalentes.
- d) Segmente a imagem com o algoritmo de crescimento de regiões, partindo do ponto central e assumindo que 2 pontos são semelhantes se a diferença absoluta entre eles for menor que 3.
- e) Calcule a circularidade, o código de cadeia, curva Phi-S e número de forma dos objetos segmentados no item (c).
- f) Determine os esqueletos dos objetos segmentados no item (c).
- g) Considerando o elemento estruturante cruz 3x3 mostre o resultado a erosão, dilatação, abertura e fechamento para a imagens binarizadas segundo os critérios do item (c).
6. Para que classes de problemas a representação por contorno é mais indicada que a representação por região? Qual as vantagens e desvantagens de se utilizar cada abordagem?

7. Calcule a Transformada de Hough para as imagens abaixo, considerando  $\theta$  como múltiplos de 45 graus e  $\rho$  variando em intervalos de 0.5 pixel.



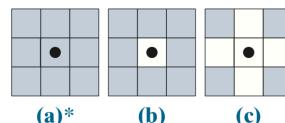
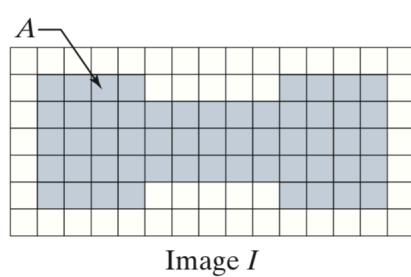
NÃO

8. A Transformada de Hough é invariante à translação, rotação e escala? Discuta.

9. Mostre o resultado

SIM

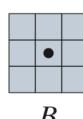
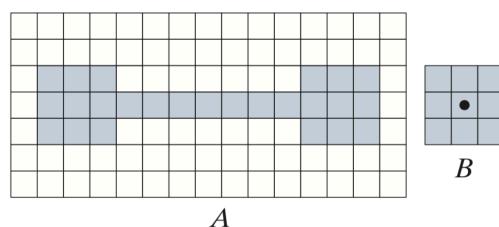
→ a posição de objetos na imagem não afeta os parâmetros detectados pela transformada (mesmo com rotação)



10. Mostre o resultado

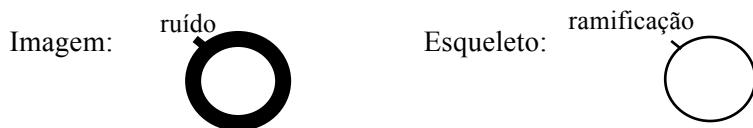
- a) da abertura do objeto A pelo elemento estruturante abaixo  
 b) do fechamento do objeto A pelo elemento estruturante abaixo

mas ela depende das dimensões de imagem e suas proporções.  
 Se mudar a escala, altera seu resultado

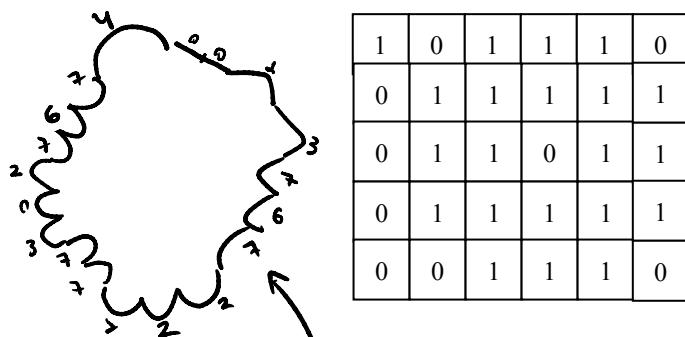


11. Repita o exercício anterior, agora considerando o elemento estruturante cruz 3x3.
12. Considere o problema de determinar o esqueleto de objetos que possuem furos, tais como ruelas. O algoritmo apresentado no livro-texto possui o inconveniente de ser sensível a ruídos nas bordas, de modo que o esqueleto pode apresentar ramificações, além do esqueleto circular que se deseja neste caso.

Ex:



Proponha um novo algoritmo, baseado no algoritmo do livro-texto, que gere um esqueleto sem ramificações, ou seja, uma curva fechada onde cada ponto tem 2 vizinhos. Determine o esqueleto da imagem abaixo, utilizando as duas versões do algoritmo.



13. Para cada número de forma abaixo, mostre a forma do objeto que representa, se possível:

$$1 \leq 3 \leq 4$$

a) 0013767221773027674

b) 11111111

c) 000200300760732002

14. Associe as imagens dos objetos *B* a *F* abaixo ao descritor que mais se aproxima das suas características de forma (contorno no sentido anti-horário). Pode haver repetições.

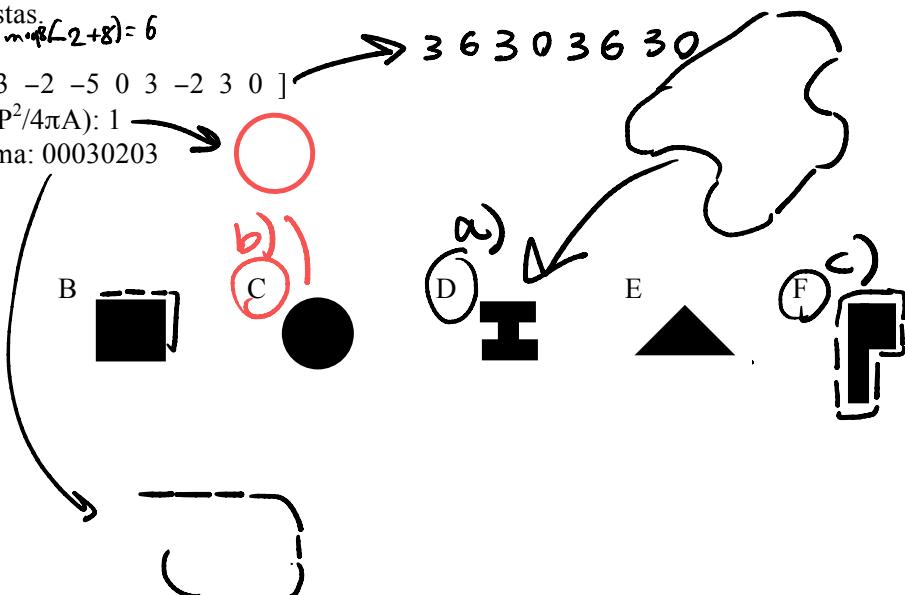
Justifique as respostas.

$$\text{máx}(2+8)=6$$

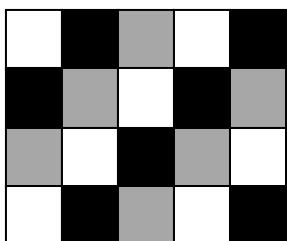
- a) curva Phi-S: [ 3 -2 -5 0 3 -2 3 0 ]
- b) compacidade ( $P^2/4\pi A$ ): 1
- c) número de forma: 00030203

A

0	0	1	1	1
1	1	1	1	1
0	1	1	1	0
0	1	1	1	1
1	0	0	1	0

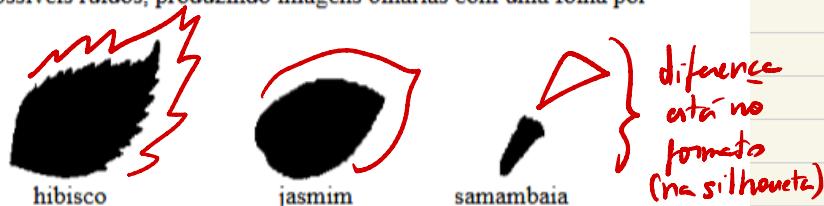


15. Dada a imagem abaixo, calcule as matrizes de cocorrência C01, C10 e C11 e respectivos descritores de entropia e homogeneidade.



$$H = \sum_{i,j} \frac{P(i,j)}{1 + |i-j|}$$

14. As imagens da figura abaixo correspondem a folhas de 3 espécies de plantas. Um pesquisador gostaria de desenvolver um sistema capaz de reconhecer a espécie a partir da imagem de uma folha, independentemente da idade da planta. Sabe-se que o formato geral da folha não se altera substancialmente com a idade, para o hibisco e o jasmim, mas a samambaia tem folhas mais curtas quando jovem. De modo geral, quanto mais jovens, menores são as folhas das plantas. As imagens são obtidas através de um scanner, que elimina possíveis ruídos, produzindo imagens binárias com uma folha por imagem.



- a) Sugira um método para segmentar a folha da imagem produzida pelo scanner, justificando sua escolha. **binarização**
  - b) Discuta as opções de representação por região e contorno e sua adequação para este problema. Indique qual a abordagem mais adequada, justificando sua opção.
  - c) Sugira um conjunto de características para descrever as folhas, de modo a facilitar o processo de classificação. Justifique a escolha destas características.
  - d) Discuta opções de classificadores para o problema. Quais informações adicionais precisariam ser obtidas do pesquisador para guiar a sua escolha?

→ KNN (atribui uma nova folha à classe com menor distância entre suas vizinhanças + proximidades) - Circunvizinhos

SVM (concentra o melhor hiperplano que maximiza a distância de reta até os support vectors).

Árvore de Decisão (tome decisão com base em características definidas de cada classe) → + interpretabilidade

relaciones más -lineales  
detrás de los grandes

15. Deseja-se construir um classificador que permita reconhecer uma ferramenta, a partir da sua imagem capturada e processada por uma câmera, classificando-a como chave de fenda (c1), chave inglesa (c2) ou martelo (c3). Para isso, foi utilizada uma amostra de 12 objetos e extraídas as características de área e retangularidade que são dadas na tabela abaixo.

Objeto	1	2	3	4	5	6	7	8	9	10	11	12
Classe	1	1	1	1	2	2	2	2	3	3	3	3
Área	20	25	20	30	80	85	80	85	80	70	70	80
Retang	0,8	0,6	0,7	0,7	0,6	0,8	0,9	0,7	0,4	0,3	0,4	0,3

- Calcule a média e a matriz de covariância para cada uma das classes
- Plote os objetos e as médias no espaço de soluções do problema
- Inverta, se possível, as 3 matrizes de covariância: A inversa  $C^{-1}$  será tal que  $C C^{-1} = I$ , o que leva ao sistema de equações

$$\begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

onde os valores de  $a, b, c$ , e  $d$  compõem a matriz de covariância inversa.

- Classifique o novo padrão  $x = [40 \ 0,5]$  utilizando os classificadores Euclídeo e de Mahalanobis. Discuta os resultados.
- Classifique o novo padrão  $x = [78 \ 0,5]$  utilizando os classificadores Euclídeo e de Mahalanobis. Discuta os resultados.

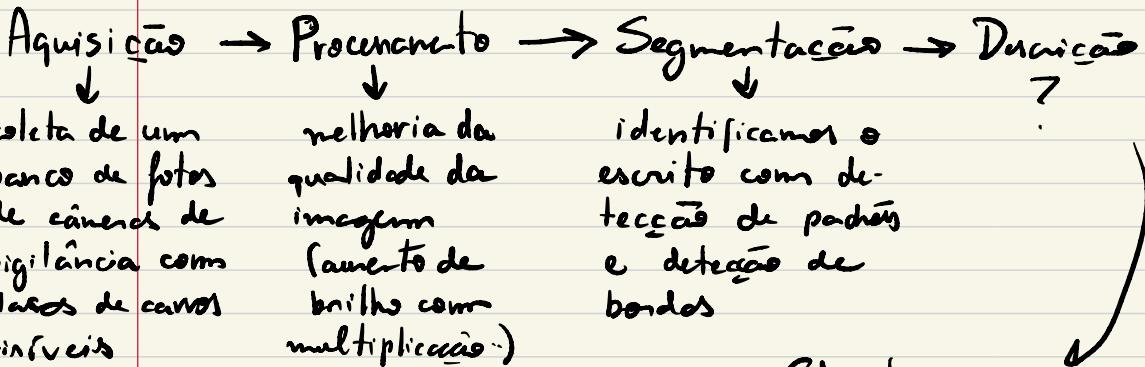
a) Médias:

$\overline{1} = \frac{20+25+20+30}{4} = 23,75$	$\overline{2} = \frac{80+85+80+85}{4} = 82,5$	$\overline{3} = \frac{80+70+70+80}{4} = 75$
$\text{retang.} = 0,7$	$= 0,75$	$= 0,35$

Matrizes de covariância  $\rightarrow (\sum (x_i - \mu_x)(y_i - \mu_y)/N-1)^T = C$

Inversa de  $C = [[a, b], [c, d]] \rightarrow (\frac{1}{ad-bc} \cdot [[d, -b], [-c, a]])$  uma matriz

1. Deseja-se projetar um sistema para leitura automática das placas de veículos em um estacionamento. Produza um roteiro para a solução do problema, caracterizando as etapas do processo de visão computacional, da aquisição ao reconhecimento. Em cada etapa, descreva uma técnica que poderia ser utilizada para a solução deste problema específico.



Classificação

treinamento das imagens p/ definir um OCR com CNN (?)

## Correções:

1- aquisição: - captura de imagens de veículos entrando e saindo de local estratégico, com placas de carros inválidos, usando câmeras de alta resolução e nitidez.

2- processamento: - preparação das imagens

- filtro de ruízização c/ mediana p/ redução de ruído
- conversão p/ escala de cinza
- ajuste de contraste com aplicações de normalização de bilha

3- segmentação: localização e detecção da região da placa com limiarização

OCR = Optical Character Recognition

CNN = Convoluted Neural Network

4- descricão: isolamento dos caracteres usando segmentação por contorno

5- classificação: reconhecimento dos caracteres usando um OCR construído com CNN

3. Na série de TV "The designated survivor" (Netflix) uma detetive tenta recuperar a impressão digital em uma taça de vinho a partir de uma foto obtida com a câmera de um smartphone. Sabe-se que os picos e vales de uma impressão digital têm em média 0,2mm de espessura. Considerando que a foto capturou um campo visual de 1 m<sup>2</sup>, qual teria que ser a resolução mínima da câmera para isso ser possível? ?

$$\text{pico/vale} = 0,2\text{mm} = 0,0002\text{m}$$

$$\text{campo visual da foto} = 1\text{m} \cdot 1\text{m} = 1\text{m}^2$$

$\rightarrow$  pixels

$$\text{nº de elementos mínimos por lado} \rightarrow \frac{1\text{m}}{0,0002\text{m}} = 5.000 \text{ pixels}$$

p/ que em detalhe apareça

$$5.000 \cdot 5.000 = 25.000.000 = 25 \text{ megapixels}$$

resolução total  
c/ os pixels necessários

se tiver menos que isso, os detalhes c/ tamanhos próximos ou menores de 0,2mm não "fundir" e não podendo ser separados, impossibilitando o reconhecimento.

## 2. Diferencie amostragem e quantização

### Amostragem

vs

### Quantização

- transformação de uma imagem contínua em uma matriz discreta de pixels.
- escolha de pontos em intervalos regulares para sua representação visual.
- basicamente, colocamos uma malha de pixels sobre a imagem e cada pixel captura o valor médio de f naquele ponto.
- define a resolução espacial da imagem.

- atribuição de valores inteiros a cada pixel, convertendo os valores contínuos de intensidade (brilho ou cor) em um nº finito de níveis digitais, determinando a profundidade tonal disponível por ponto.
- basicamente, buscamos uma cor média que consiga representar bem as cores de um espaço delimitado.

Amostragem é a escolha de um ponto dentro de um espaço, ou melhor, a escolha de um conjunto de pontos dentro de um espaço real que irão compor a imagem digital.

Definir a amostragem é conseguir selecionar uma quantidade finita de pontos de uma imagem com infinitos pontos, em outras palavras, é conseguir selecionar pequenos conjuntos de espaços de forma que, devidamente arranjados, estes conjuntos se tornem a imagem discreta.

A amostragem está diretamente ligada com a quantidade de informação que se deseja guardar. Quanto maior a amostragem, mais detalhes teremos e consequentemente maior será o espaço necessário para o armazenamento.

### Quantização

É a escolha de uma cor dentro de uma imagem, o processo de quantização leva em conta cada uma das cores dentro de uma sub-imagem e escolhe uma. Existem diversas técnicas de quantização de cor dentro de um espaço, basicamente o que estas técnicas fazem é buscar uma cor media que consiga representar bem as cores de um espaço delimitado.

Processo	O que é discretizado?	Impacto na imagem
Amostragem	Espaço (posição dos pixels)	Resolução (detalhamento)
Quantização	Intensidade/amplitude	Profundidade tonal (quantidade de cores/tons)

4. Porque o problema da segmentação é considerado sem solução geral?

O problema da segmentação de imagens é considerado sem solução geral pois existem diversas formas de resolvê-lo, não somente uma, já que cada imagem possui características distintas como texturas, iluminação, ruído, cor e sobreposição de objetos, que determinarão qual é a melhor abordagem a ser levada.

No entanto, ela não tem solução geral definida pois é ambígua no que consideramos como uma divisão de regiões correta, depende do contexto, das propriedades da imagem e do objetivo final, adaptando técnicas (que podem não ser perfeitas) para cada caso.

5. Para cada imagem abaixo:

A	3	5	2	1	1
	1	4	6	2	1
	1	1	5	6	2
	1	1	1	1	1
	1	2	2	2	1

B	5	1	2	1	8
	6	6	5	6	1
	2	1	8	7	7
	6	1	2	8	8
	7	8	2	1	1

C	1	1	9	1	1
	1	1	9	8	7
	9	9	9	2	1
	1	1	2	8	8
	1	2	2	8	9

- Determine sua resolução e quantização.
- Segmenta a imagem pela técnica de limiar, com valor de corte  $t > 2$  para A,  $t > 5$  para B e  $t > 6$  para C.
- Aplique o algoritmo de rotulação no resultado do item anterior. Diga como fica a tabela de rótulos equivalentes.
- Segmenta a imagem com o algoritmo de crescimento de regiões, partindo do ponto central e assumindo que 2 pontos são semelhantes se a diferença absoluta entre eles for menor que 3.
- Calcule a circularidade, o código de cadeia, curva Phi-S e número de forma dos objetos segmentados no item (c).
- Determine os esqueletos dos objetos segmentados no item (c).
- Considerando o elemento estruturante cruz  $3 \times 3$  mostre o resultado a erosão, dilatação, abertura e fechamento para a imagens binarizadas segundo os critérios do item (c).

5-a) Resolução de todas as imagens:  $5 \times 5 = 25$

A	3	5	2	1	1
	1	4	6	2	1
	1	1	5	6	2
	1	1	1	1	1
	1	2	2	2	1

B	5	1	2	1	8
	6	6	5	6	1
	2	1	8	7	7
	6	N	2	8	8
	7	8	X	1	1

C	1	1	9	1	1
	1	X	9	8	7
	9	9	9	2	1
	1	1	X	8	8
	1	2	2	8	9

Quant. = 6 (1 a 6)  
1, 2, 3, 4, 5, 6  
( $\hookrightarrow$  6 níveis distintos)

$Q = 6$   
1, 2, 5, 6, 7, 8

$Q = 5$   
1, 2, 7, 8, 9

b)  $+ > 2$

$+ > 5$

$+ > 6$

A	1	1	2	1	1
	1	1	2	1	1
	1	1	1	2	1
	1	1	1	1	1
	1	2	2	2	1

B	5	1	2	1	8
	6	6	5	6	1
	2	1	8	7	7
	6	N	2	8	8
	7	8	X	1	1

C	1	1	9	1	1
	1	1	9	8	7
	9	9	9	2	1
	1	1	8	8	8
	1	2	2	8	9

1	1	0	0	9
0	1	1	0	0
0	0	1	1	0
9	0	0	0	0
0	0	0	0	0

0	0	0	0	1
1	1	0	1	0
0	0	1	1	2
1	0	0	1	1
1	1	0	0	0

0	0	1	0	0
0	0	1	1	1
1	1	1	0	0
0	0	0	1	1
0	0	0	1	1

c)	1	+	0	0	9
	0	+	1	0	0
	0	0	+	1	0
	9	0	0	0	0
	0	0	0	0	0

0	0	0	0	1
1	1	0	1	0
0	0	1	1	1
1	0	0	1	1
1	1	0	0	0

0	0	1	0	0
0	0	1	1	1
1	1	1	0	0
0	0	0	1	1
0	0	0	1	1

d)

A	3	5	2	1	1
	1	4	6	2	1
	1	1	5	6	2
	1	1	1	1	1
	1	2	2	2	1

B	5	1	2	1	8
	6	6	5	6	1
	2	1	8	7	7
	6	1	2	8	8
	7	8	2	1	1

C	1	1	9	1	1
	1	1	9	8	7
	9	9	9	2	1
	1	1	2	8	8
	1	2	2	8	9

\* **crescimento de regiões**  $\Rightarrow$  partindo do ponto central, adiciona-se ao grupo qualquer pixel vizinho cujo valor tenha diferença absoluta menor que 3 em relação ao valor do centro atual.

e)

A	3	5	2	1	1
	1	4	6	2	1
	1	1	5	6	2
	1	1	1	1	1
	1	2	2	2	1

B	5	1	2	1	8
	6	6	5	6	1
	2	1	8	7	7
	6	1	2	8	8
	7	8	2	1	1

C	1	1	9	1	1
	1	1	9	8	7
	9	9	9	2	1
	1	1	2	8	8
	1	2	2	8	9

\* **circularidade**  $\Rightarrow$  mede o quanto a forma está de um círculo. Fórmula:

$$C = \frac{4\pi \cdot \text{Área}}{\text{Perímetro}^2}$$

n° de pixels do  
objeto

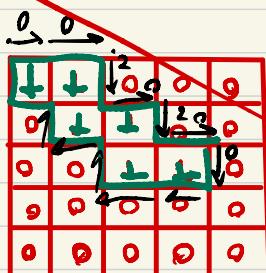
n° de pixels na  
borda do objeto

→ **código de cadeia (chain code)**  $\Rightarrow$  representa a borda do objeto como uma sequência de direções

- conectando pelo pixel da borda superior, siga ao redor do contorno do objeto e registre direção entre pixels adjacentes
- 0 = direita, 1 = diagonal inferior direita, 2 = baixo, 3 = diagonal inferior esquerda, 4 = esquerda, 5 = diagonal superior esquerda, 6 = cima, 7 = diagonal superior direita.

\* curva phi-S  $\Rightarrow$  gráfico que relaciona o comprimento do contorno (perímetro) c/ a soma dos ângulos entre vetores consecutivos da borda

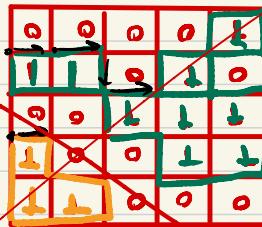
$$\text{nº de forma} = \frac{\text{Perímetro}}{\text{Área}}$$



$$* C = \frac{4\pi \cdot 6}{6^2} = \frac{24\pi}{36}$$

\* 0,0,2,0,2,2,0,  
4,4,6,4,6,4,6.

$$* 6/6 = 1$$



$$* C = \frac{4\pi \cdot 9}{9^2} = \frac{36\pi}{81}$$

$$C = \frac{4\pi \cdot 3}{3^2} = \frac{12\pi}{9}$$

\* 0,0,2,0,6,0,6,0,  
2,4,2,0,2,2,4,4,  
6,4,6,4,4,6.

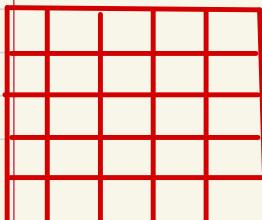
$$* C = \frac{4\pi \cdot 11}{11^2} = \frac{44\pi}{121}$$

\* 0,0,6,6,0,2,0,0,  
2,4,4,2,0,0,2,2,  
4,4,6,5,4,4,4,6.

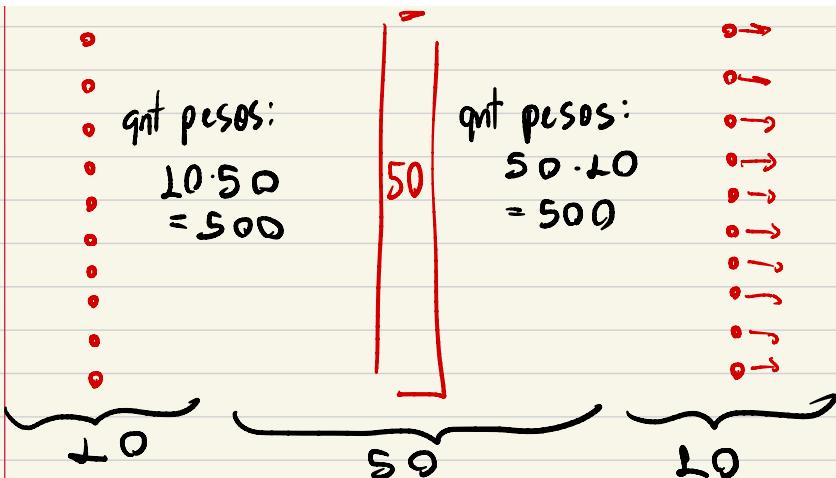
$$* 11/11 = 1$$

0,2,0,2,4,4,6,6.

$$* 3/3 = 1$$



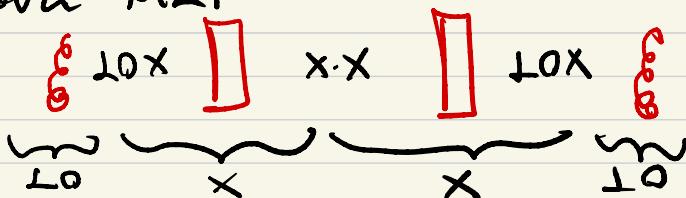
1. Uma MLP possui 10 unidades de entrada, 50 unidades na camada escondida e 10 unidades na camada de saída (sem contar o bias). Deseja-se substituir a camada escondida por 2, cada uma com  $n$  unidades, sem aumentar o número total de pesos da rede original. Qual o valor máximo de  $n$ ?



$$n^{\circ} \text{ de pesos entre 2 camadas} = n^{\circ} \text{ de neurônios da 1^{\circ} camada} \cdot n^{\circ} \text{ de neurônios da 2^{\circ} camada}$$

$$\begin{aligned} \text{total} &= 500 + 500 \\ \text{de pesos} &= 1000 \end{aligned}$$

nova MLP:



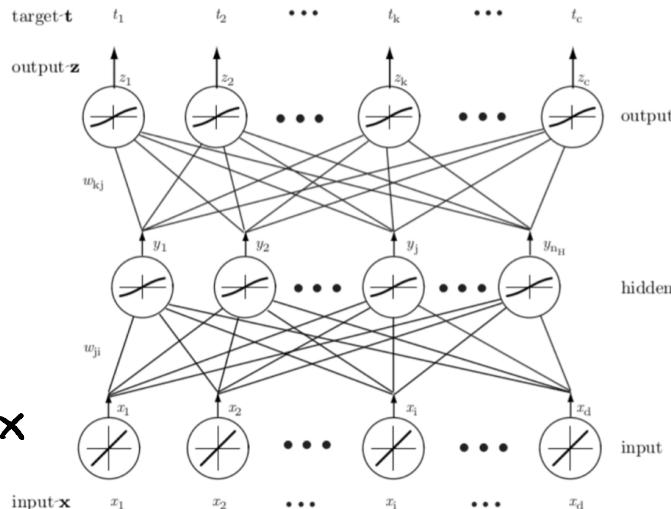
$$10 \cdot x + x \cdot x + 10 \cdot x = 1000$$

$$x^2 + 2(10x) = 1000$$

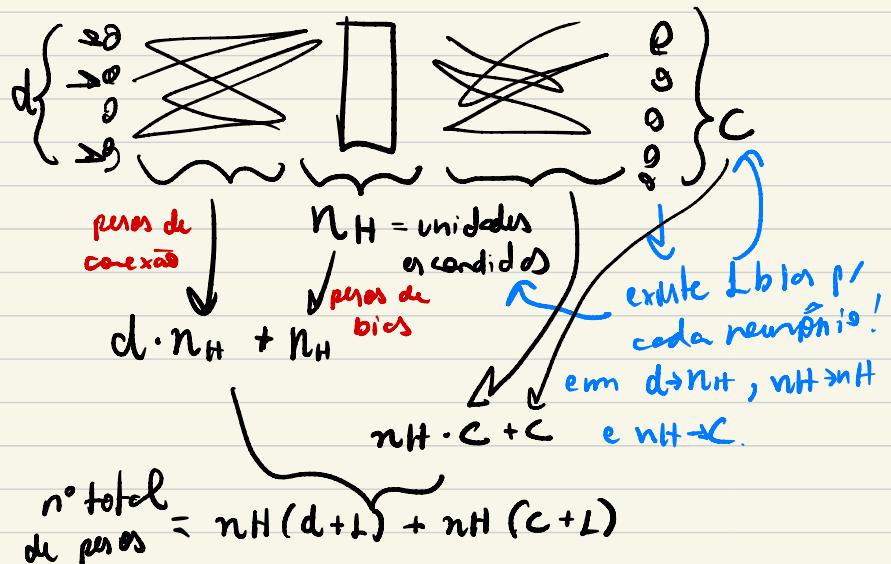
$$x^2 + 20x - 1000 = 0$$

$$= 23,135 \rightarrow \approx 23$$

2. Considere uma rede padrão de 3 camadas cuja entrada  $x$  possui dimensão  $d \times 1$ , a primeira camada da rede possui  $d$  unidades de entrada e possui somente uma ativação linear do tipo  $f(x) = x$ , a camada escondida possui  $n_H$  unidades escondidas e a camada final possui  $c$  unidades de saída e o bias. Qual o número total de pesos que existem na rede?

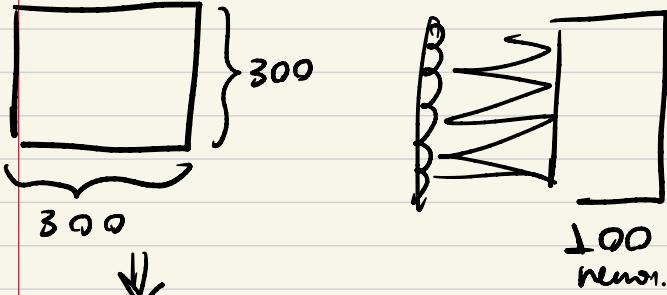


$$\text{entrada} = d \times 1$$



3. Dado uma imagem de  $300 \times 300$  pixels colorida (RGB) como entrada para alguns modelos, responda às questões abaixo.

a) Modelo 1: Suponha que você não esteja usando uma rede convolucional. Se a primeira camada oculta tiver 100 neurônios, cada um deles totalmente conectado à entrada, quantos parâmetros essa camada oculta possui (incluindo os parâmetros do bias)?



$90.000 \text{ pixels}$

cada pixel RGB  
tem 3 canais

$$90.000 \cdot 3 = 270.000$$

$$\begin{aligned} & 100 \cdot 270.000 \\ & = 27.000.000 \text{ pesos de conexões} \\ & (\text{cada unidade tem } 1 \text{ peso}) + 100 \text{ bias} \end{aligned}$$

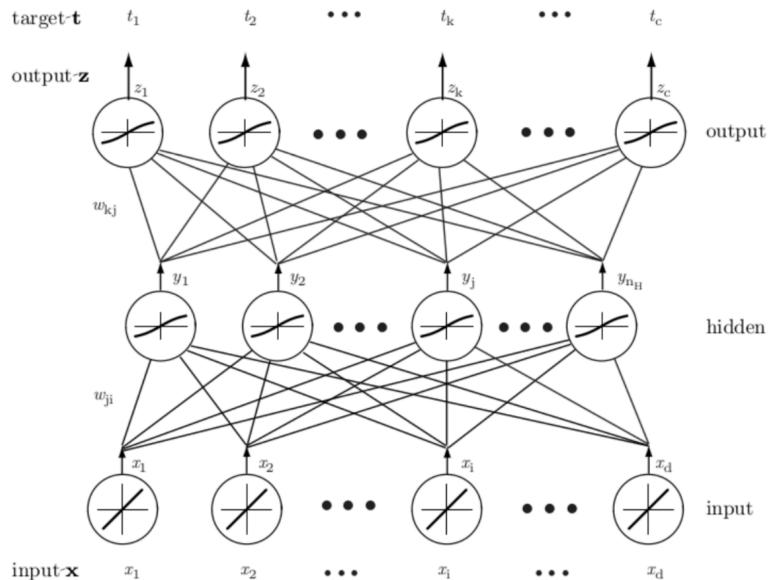
↓  
*(de cada neurônio)*

$$27.000.000 + 100$$

$$= \underline{\underline{27000100}}$$

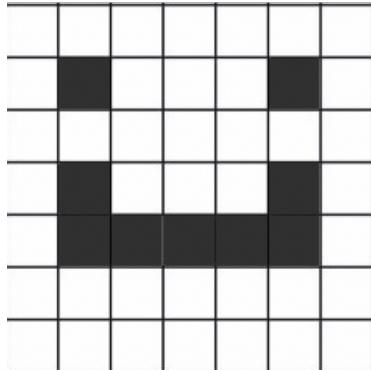
### Lista de Exercícios No. 2

- Uma MLP possui 10 unidades de entrada, 50 unidades na camada escondida e 10 unidades na camada de saída (sem contar o bias). Deseja-se substituir a camada escondida por 2, cada uma com  $n$  unidades, sem aumentar o número total de pesos da rede original. Qual o valor máximo de  $n$ ?
- Considere uma rede padrão de 3 camadas cuja entrada  $x$  possui dimensão  $d \times 1$ , a primeira camada da rede possui  $d$  unidades de entrada e possui somente uma ativação linear do tipo  $f(x) = x$ , a camada escondida possui  $n_H$  unidades escondidas e a camada final possui  $c$  unidades de saída e o bias. Qual o número total de pesos que existem na rede?



- Dado uma imagem de  $300 \times 300$  pixels colorida (RGB) como entrada para alguns modelos, responda às questões abaixo.
  - Modelo 1: Suponha que você não esteja usando uma rede convolucional. Se a primeira camada oculta tiver 100 neurônios, cada um deles totalmente conectado à entrada, quantos parâmetros essa camada oculta possui (incluindo os parâmetros do bias)?
  - Modelo 2: Suponha agora que você use uma camada convolucional com 100 filtros de  $5 \times 5$  cada. Quantos parâmetros essa camada oculta possui (incluindo os parâmetros de bias)?

4. Dado a imagem  $7 \times 7$  abaixo, aplique um filtro que seja capaz de detectar somente os olhos dessa representação (quase perfeita) do rosto humano. Você deve pensar nos valores e no tamanho do filtro que irá utilizar, além de aplicá-lo à imagem e mostrar o resultado obtido. Assuma que os pixels brancos possuem valor igual a 0 e os pixels pretos possuem valor igual a 1.



5. Dado uma imagem em preto e branco de tamanho  $8 \times 8$  pixels e um filtro de tamanho  $3 \times 3$ , indique as dimensões da matriz resultante da convolução e o tamanho do padding que deverá ser utilizado em cada um dos casos:
- Valid padding
  - Same padding
6. Suponha uma entrada de tamanho  $63 \times 63 \times 16$ . Ao aplicar uma convolução nessa entrada com 32 filtros de tamanho  $7 \times 7$ , usando stride igual a 2 e sem padding. Qual será o volume de saída?
7. Suponha uma entrada de tamanho  $15 \times 15 \times 8$ . Usando a operação de padding com  $p=2$ , qual é a dimensão do dado de saída após o padding?
8. Dado uma entrada de dimensão  $63 \times 63 \times 16$  e uma convolução com 32 filtros de dimensão  $7 \times 7$  cada e um stride igual a 1, qual deverá ser o tamanho do padding utilizado para que você obtenha uma saída com o mesmo tamanho da entrada (same padding)?

9. Considere um volume de entrada  $65 \times 65 \times 3$  e um filtro  $11 \times 11 \times 3$ . Quantas operações de multiplicação serão feitas em cada um dos casos:
  - a) Valid padding e stride = 1
  - b) Valid padding e stride = 3
  - c) Same padding e stride = 1
  - d) Same padding e stride = 3
10. Suponha uma entrada de tamanho  $32 \times 32 \times 16$ . Seja a aplicação do max pooling com stride e tamanho de filtro iguais a 2. Quais são as dimensões da saída?
11. Suponha uma entrada de tamanho  $6 \times 6 \times 3$ . Seja a aplicação de um pooling (average ou max) com stride e tamanho de filtro iguais a 2. Responda:
  - a) Quais são as dimensões da saída?
  - b) Assumindo que os valores do primeiro canal estão mostrados na matriz abaixo, mostre o resultado obtido ao aplicar o seguinte Max pooling e Average pooling

$[[4\ 9\ 2\ 5\ 8\ 3]\ [5\ 6\ 2\ 4\ 0\ 3]\ [2\ 4\ 5\ 4\ 5\ 2]\ [5\ 6\ 5\ 4\ 7\ 8]\ [5\ 7\ 7\ 9\ 2\ 1]\ [5\ 8\ 5\ 3\ 8\ 4]]$
12. Suponha que a entrada para uma rede neural de convolução seja uma imagem colorida (RGB)  $32 \times 32$ . A primeira camada contém oito filtros  $5 \times 5$  com três canais, utilizando Valid padding e stride = 2. Qual o formato da saída dessa camada?
13. Dado uma imagem de dimensão  $224 \times 224$  com 3 canais (RGB), desenhe a rede convolucional, incluindo as dimensões das matrizes de entrada e saída, de acordo com as operações descritas abaixo.
  - a) Aplique uma convolução com "Valid padding" com 96 filtros de tamanho 7 e stride igual a 2. Em seguida, aplique um max pooling com filtro de tamanho 3 e stride igual a 2. A saída dessa camada será chamada de  $A^{[1]}$ .
  - b) Aplique uma convolução com "Valid padding" com 256 filtros de tamanho 5 e stride igual a 2. Em seguida, aplique um max pooling com filtro de tamanho 3 e stride igual a 2. A saída dessa camada será chamada de  $A^{[2]}$ .
  - c) Aplique uma convolução com "Same padding" com 384 filtros de tamanho 3 e stride igual a 1. A saída dessa camada será chamada de  $A^{[3]}$ .
  - d) Aplique uma convolução com "Same padding" com 384 filtros de tamanho 3 e stride igual a 1. A saída dessa camada será chamada de  $A^{[4]}$ .

- e) Aplique uma convolução com "Same padding" com 256 filtros de tamanho 3 e stride igual a 1. Em seguida, aplique um max pooling com filtro de tamanho 3 e stride igual a 2. A saída dessa camada será chamada de  $A^{[5]}$ .
- f) Aplique uma camada fully-connected com 4096 nós. A saída dessa camada será chamada de  $A^{[6]}$ .
- g) Aplique uma camada fully-connected com 4096 nós. A saída dessa camada será chamada de  $A^{[7]}$ .
- h) Por fim, aplique uma softmax (aqui não é necessário se preocupar com a dimensão da saída). A saída dessa camada será chamada de  $A^{[8]}$ .

14. Dada a função  $f(x)=\cos \omega x$ , definida no intervalo  $0 \leq x < 2$ :

- a) Gere um vetor contendo a amostragem da função nos pontos  $x=0$  e  $x=1$ .
- b) Calcule a DFT sobre o vetor
- c) Baseado nos coeficientes encontrados, desenhe os componentes da série e a função reconstituída.

15. Faça o mesmo procedimento do item anterior para a função  $f(x)=\cos 2\omega x$ , definida no intervalo  $0 \leq x < 4$  e amostrada nos pontos  $x=0, 1, 2$  e  $3$ . Compare os resultados obtidos.

16. Dados os espectros de Fourier abaixo, determine a imagem correspondente.

A	3.0	-0.5 + 0.69i	-0.5 + 0.16i	-0.5 - 0.16i	-0.5 - 0.69i
---	-----	-----------------	-----------------	-----------------	-----------------

B	1.5	-0.25 - 0.25i	0	-0.25 + 0.25i
---	-----	------------------	---	------------------

17. Dadas as imagens abaixo, calcule a DFT correspondente. Compare as imagens e comente os resultados à luz da propriedade da translação.

A	0	1	0	0	0
---	---	---	---	---	---

B	0	0	1	0	0
---	---	---	---	---	---

18. Dadas as imagens abaixo, considerando pontos externos como possuindo valor 0:

A	1	2	0	2	1
---	---	---	---	---	---

B	3	2	1	2	3
---	---	---	---	---	---

C	1	2		
	2	3		

D	1	2	1		
	2	8	2		
	1	2	1		

- a) Calcule  $\mathbf{A} * \mathbf{B}$
- b) Calcule  $\mathbf{B} * \mathbf{A}$
- c) Calcule  $\mathbf{C} * \mathbf{D}$
- d) Calcule  $\mathbf{D} * \mathbf{C}$
- e) Calcule a DFT para as imagens **A** e **B**. Calcule a DFT inversa sobre o resultado.

- f) Aplique filtros passa-baixa nas imagens **A** e **B** com frequência de corte  $|u|<2$ .
- g) Aplique filtros passa-alta nas imagens **A** e **B** com frequência de corte  $|u|>1$ .

19. Para cada imagem abaixo, considerando pontos externos como indefinidos:

**A**

3	5	2	1	1
1	4	6	2	1
1	1	5	6	2
1	1	1	1	1
1	2	2	2	1

**B**

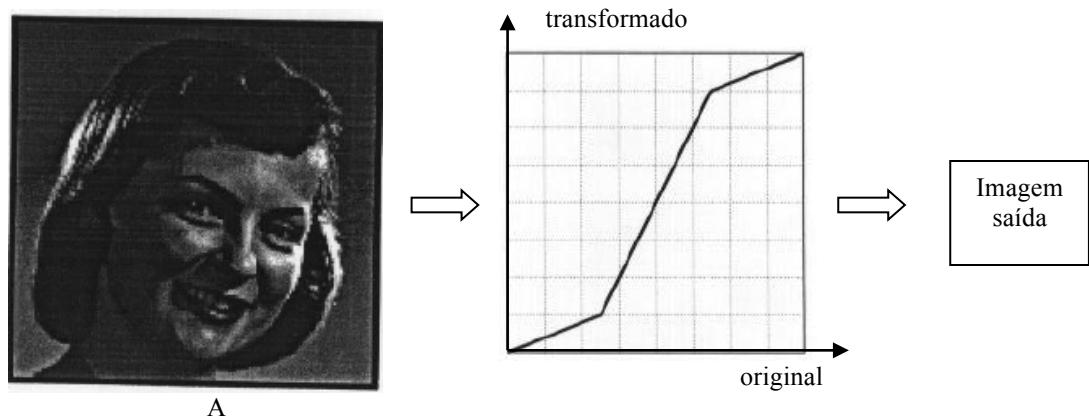
5	1	2	1	8
6	6	5	6	1
2	1	8	7	7
6	1	2	8	8
7	8	2	1	1

**C**

1	1	9	1	1
1	1	9	8	7
9	9	9	2	1
1	1	2	8	8
1	2	2	8	9

- a) Determine o histograma de freqüências
- b) Aplique um filtro de suavização 3x3 pela média
- c) Aplique um filtro de suavização 3x3 pela mediana
- d) Altere o contraste da imagem através da equalização do histograma. As novas intensidades devem variar entre 0 e 255.
- e) Realce as bordas da imagem, através de filtros de Sobel.

20. O gráfico abaixo representa a função de transformação de histograma aplicada à imagem A.

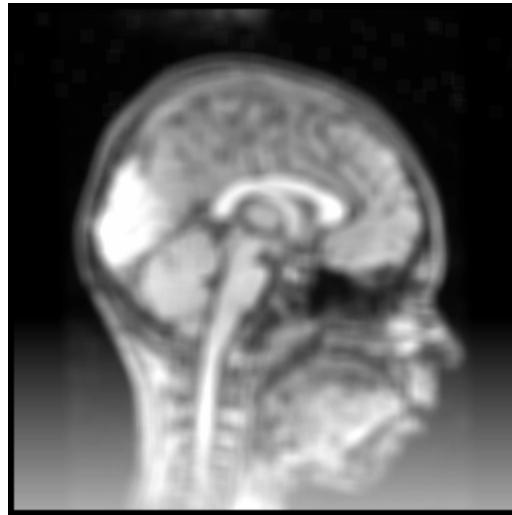


- a) Caracterize a imagem de saída quanto ao seu tamanho e conteúdo.
- b) Para que são usadas as funções de transformação de histograma?
- c) É possível aplicar uma transformação de histograma na qual 2 pixels de tons de cinza diferentes da imagem de entrada passem a ter o mesmo valor após a transformação? Justifique.

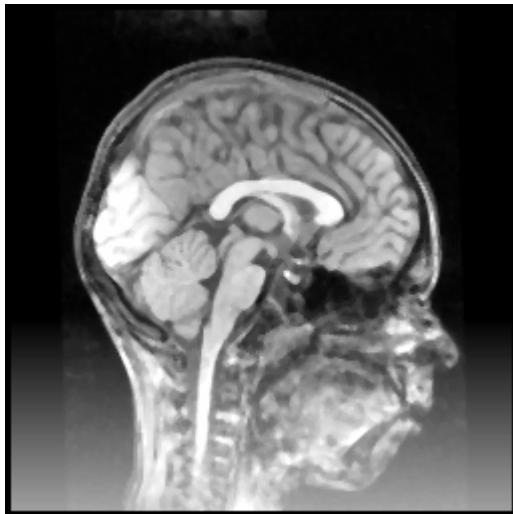
21. Considere a imagem original A e as imagens B, C e D obtidas a partir de A:



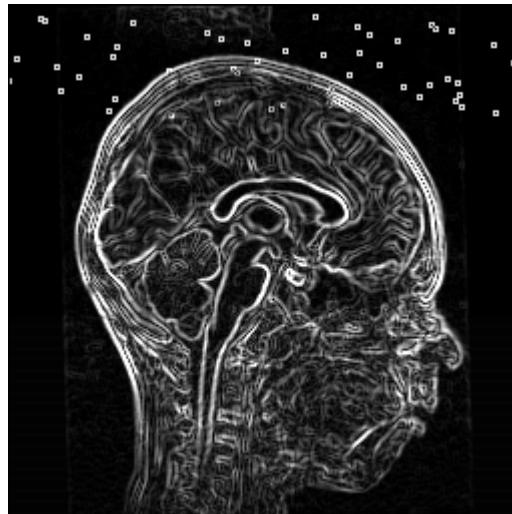
A



B



C



D

- a) Indique os elementos de baixa freqüência presentes na imagem original A.
- b) Indique os elementos de alta freqüência presentes na imagem original A.
- c) Descreva o processo aplicado a A para se obter B. Justifique a resposta.
- d) Descreva o processo aplicado a A para se obter C. Justifique a resposta.
- e) Descreva o processo aplicado a A para se obter D. Justifique a resposta.