

### Lista de Exercícios No. 1

1. Caracterize brevemente os passos do processo de visão computacional.
2. Diferencie amostragem e quantização
3. Porque o problema da segmentação é considerado sem solução geral?
4. 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

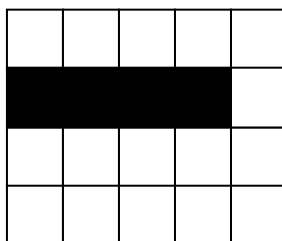
B

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

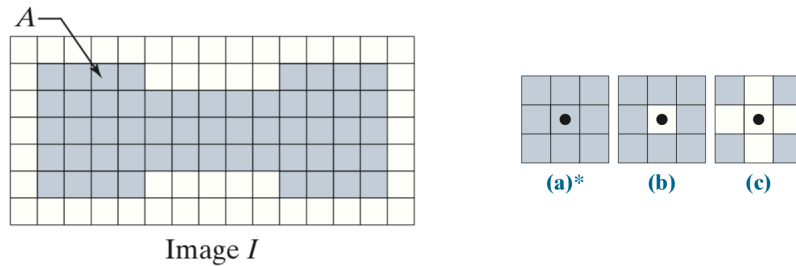
C

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

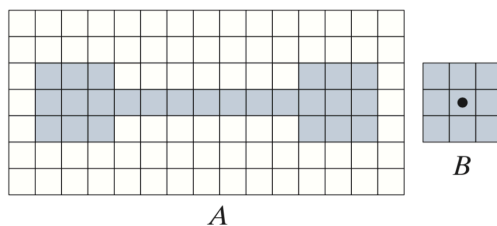
- a) Segmente a imagem pela técnica de limiar, com valor de corte  $t \leq 2$  para A,  $t \leq 4$  para B e  $t \leq 6$  para C.
  - b) Aplique o algoritmo de rotulação no resultado do item anterior. Diga como fica a tabela de rótulos equivalentes.
  - c) Segmente a imagem com o algoritmo de crescimento de regiões, partindo do ponto central e assumindo que 2 pontos são semelhantes se a diferença absoluta entre eles for menor que 3.
  - d) Calcule a circularidade, o código de cadeia, curva Phi-S e número de forma dos objetos segmentados no item (c).
  - e) Determine os esqueletos dos objetos segmentados no item (c).
  - f) Segmente a imagem usando o algoritmo de Otsu.
5. 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?
  6. Calcule a Transformada de Hough para a imagem abaixo, considerando  $\theta$  como múltiplos de 45 graus e  $\rho$  variando em intervalos de 0.5 pixel.



7. A Transformada de Hough é invariante à translação, rotação e escala? Discuta.
8. Mostre o resultado
- da erosão da imagem  $I$  por cada um dos elementos estruturantes abaixo
  - da dilatação da imagem  $I$  por cada um dos elementos estruturantes abaixo

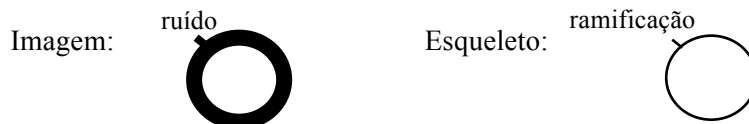


9. Mostre o resultado
- da abertura do objeto A pelo elemento estruturante abaixo
  - do fechamento do objeto A pelo elemento estruturante abaixo



10. 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.

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

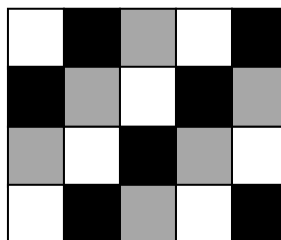
11. Para cada número de forma abaixo, mostre a forma do objeto que representa, se possível:
- a) 0013767221773027674
  - b) 11111111
  - c) 000200300760732002
12. Associe as imagens dos objetos *B* a *F* abaixo ao descritor que mais se aproxima das suas características de forma (contorno no sentido anti-horário). Pode haver repetições. Justifique as respostas.
- 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

B  C  D  E  F 

13. Dada a imagem abaixo, calcule as matrizes de coocorrência *C01*, *C10* e *C11* e respectivos descritores de entropia e homogeneidade.



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.



hibisco



jasmim



samambaia

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

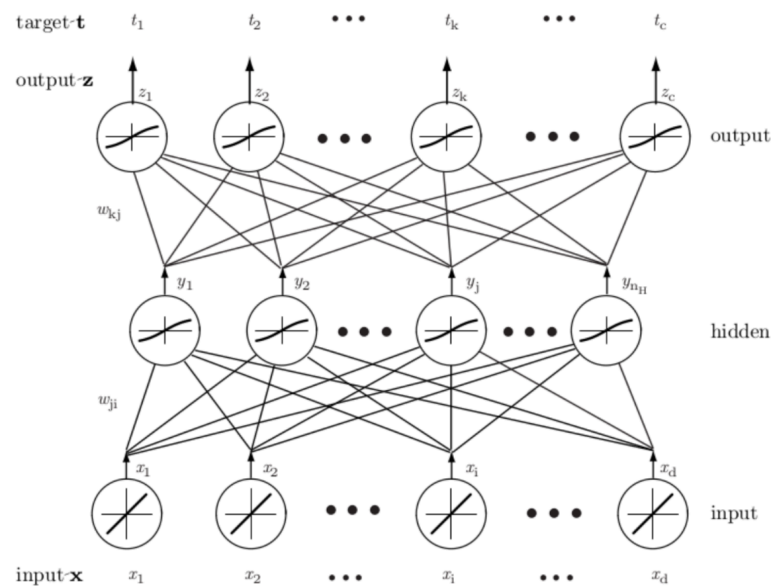
- Calcule a média e a matriz de covariância para cada uma das classes
- Plote os objetos e as médias no espaço de soluções do problema
- Inverta, se possível, as 3 matrizes de covariância: A inversa  $C^{-1}$  será tal que  $C C^{-1} = I$ , o que leva ao sistema de equações

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

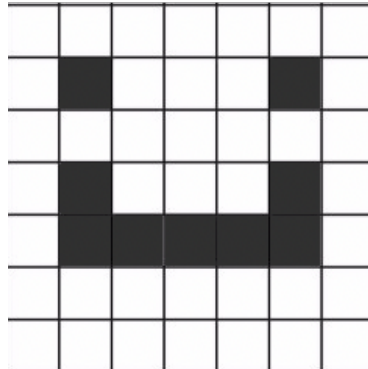
onde os valores de  $a, b, c$ , e  $d$  compõem a matriz de covariância inversa.

- Classifique o novo padrão  $x = [40 \ 0.5]$  utilizando os classificadores Euclidiano e de Mahalanobis. Discuta os resultados.
- Classifique o novo padrão  $x = [78 \ 0.5]$  utilizando os classificadores Euclidiano e de Mahalanobis. Discuta os resultados.

16. 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  $n$ ?
17. 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?



18. 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)?
  - b) 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)?
19. 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.



20. 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
21. 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?
22. 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?
23. 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)?
24. 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:
- Valid padding e stride = 1
  - Valid padding e stride = 3
  - Same padding e stride = 1
  - Same padding e stride = 3
25. 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?
26. 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]]

27. 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?
28. 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 n'os. A saída dessa camada será chamada de  $A^{[6]}$ .
  - g) Aplique uma camada fully-connected com 4096 n'os. 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  $A^{[8]}$ .