

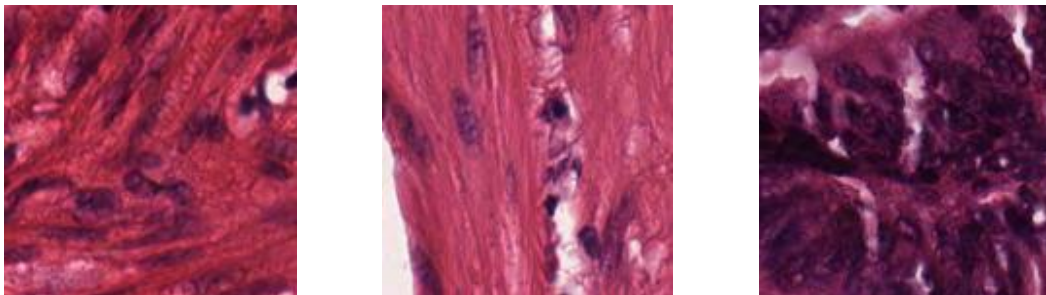
Trabalho 2

Instruções

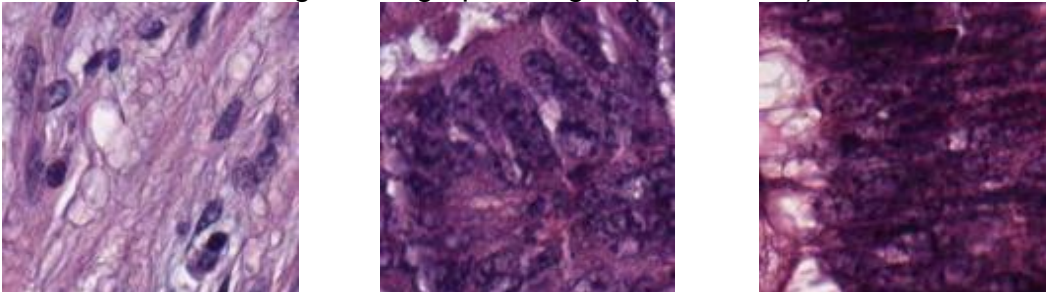
Este trabalho consistirá em uma abordagem prática da disciplina de Processamento de Imagens Digitais.

Será utilizado parte de um conjunto de imagens histológicas publicamente disponível (<https://www.nature.com/articles/srep27988#Sec2>). Nessa base imagens, foram extraídas biópsias de pacientes com e sem câncer colorretal (Figura 1). A partir da biópsia as imagens foram geradas, tirando fotos a partir de microscopia e um processo de tingimento dos tecidos para destacar as estruturas presentes na biópsia.

Figura 1: Exemplo do conjunto de dados para o grupo benigno e maligno



Imagens do grupo benigno (sem câncer)



Imagens do grupo maligno (com câncer)

Objetivos

- Realizar a extração de características de textura nos conjuntos de imagens de cada grupo
- A partir das características extraídas realizar a classificação. Nessa etapa, os atributos extraídos são fornecidos como entrada para algoritmos de *machine learning*. Se as características forem bastante diferentes entre os grupos com e sem a doença, os algoritmos de classificação conseguirão fazer a distinção dos grupos com altas taxas de acurácia.

Especificações

O trabalho poderá ser feito individualmente, em dupla ou trio

Linguagens recomendadas: matlab, octave, scilab ou python

(Caso queira usar uma linguagem não listada, favor informar)

Um relatório deverá ser feito sobre os resultados obtidos, analisando as características e os resultados obtidos na classificação, indicando quais as características mais relevantes e os melhores métodos de classificação

Os códigos também deverão ser incluídos na entrega.

Filtragem

Filtros passa-alta ou passa-baixa podem ser aplicados para analisar os resultados da extração e por consequência a classificação das imagens, verificando se contribuíram para uma maior diferenciação dos conjuntos.

Observação: A etapa de filtragem é OPCIONAL, ficando a critério da equipe o desenvolvimento ou não dessa etapa

Extração

Implementar (sem funções pré-definidas) os algoritmos para extrair descritores de textura estatísticos a partir do histograma de intensidade de uma imagem: média, desvio padrão, terceiro momento (assimetria do histograma), uniformidade e entropia.

Escolher um algoritmo de extração de características da escolha da equipe para compor o vetor de características de cada imagem

Observação: realizar a extração das características em níveis de cinza e para cada canal separadamente (R, G e B)

O vetor de características para cada imagem consistirá em 4 “tabelas” (1 em níveis de cinza e as 3 para cada canal) conforme representado. Na etapa de classificação uma última coluna será acrescentada que consistirá no rótulo daquela imagem.

Nome	Média	Desvio	3º Mom	Uniform	Entropia	Carac Grupo
Img 1						
Img 2						
...						
Img 40						

Classificação

Realizar a classificação dos dois grupos de imagens a partir das características extraídas, para tal, podem ser utilizados pacotes/funções/métodos de *machine learning* presentes nas próprias linguagens como Matlab ou Python (*scikit learning*).

Outra opção para realizar a classificação, consiste no uso do programa Weka (<https://www.cs.waikato.ac.nz/ml/weka/>), que fornece uma interface gráfica para classificação dos conjuntos, utilizando um arquivo de configuração .arff que especifica as classes e os atributos a serem classificados.

Considerar os algoritmos:

- Decision Tree (também conhecido como J48)
- Support Vector Machine (também conhecido como SMO)
- Redes Neurais (MultiLayer Perceptron)