

Relatório do trabalho prático – Classificação de imagens de cintilografia de tireoide

INTRODUÇÃO

Este relatório refere-se a experimentos práticos do trabalho de classificação de imagens de cintilografia de tireoide, proposto na disciplina de Processamento de Imagens Biomédicas. Foi fornecido um arquivo zipado com imagens no formato DICOM (*Digital Imaging and Communications in Medicine*) anonimizadas de 16-bits, com 128 x 128 pixels de dimensão, de 12 pacientes - sendo seis delas de pacientes com diagnóstico de bócio multinodular tóxico (BMT) e seis de pacientes com diagnóstico de doença de Graves.

A proposta do trabalho solicitava que fosse desenvolvido um algoritmo de visão computacional para classificar a qual das duas classes a imagem do exame de cintilografia de tireoide de um determinado paciente pertence. Foi solicitado ainda o uso da estratégia "*leave-one-patient-out*" para fazer os testes. Tal técnica consiste em treinar o modelo do classificador com o número total de pacientes -1 e testá-lo com o paciente que ficou de fora, alternando o paciente do teste até completar a classificação de toda a amostra. Como a amostra é pequena, a especificação alertava que não seria adequado o uso de redes neurais convolucionais.

A doença de Graves é um distúrbio autoimune e a principal causa de hipertireoidismo, correspondendo a mais de 70% dos casos¹. Os anticorpos desses pacientes estimulam a tireoide a produzir e secretar os hormônios tireoidianos em excesso no sangue, levando quase sempre a um aumento da glândula, chamado de bócio². Na Doença de Graves, há um quadro de bócio difuso hipercaptante: a glândula inteira - e não apenas uma área da tireoide - está trabalhando muito de forma homogênea, ficando aumentada como um todo².

Já o bócio multinodular tóxico (BMT) ou doença de Plummer caracteriza-se pela presença de múltiplos nódulos - pequenos caroços - na tireoide, sendo que um ou mais desses nódulos podem começar a produzir e secretar hormônio tireoidiano em excesso.

CARACTERÍSTICAS DA MÁQUINA, VERSÃO DO PYTHON E DETALHES GERAIS DA IMPLEMENTAÇÃO

Quanto às características da máquina, todos os experimentos foram realizados em um *notebook* Asus com arquitetura x86_64, com ordem de bytes *little endian* e processador Intel® Core™ i7-8550U CPU @ 1.80GHz, com 8 núcleos virtuais (2 por núcleo físico). Sobre o sistema operacional, a máquina operou no Ubuntu 20.04.4 LTS, versão do GNOME 3.36.8. Versão do Python utilizada: Python 3.8.10.

Na implementação, foram utilizadas funções das bibliotecas `os`, `sys`, `numpy`, `matplotlib`, `pydicom`, `opencv`, `scikit-image` e `scikit-learn` na implementação. Optou-se por modularizar o algoritmo em dois códigos:

- `thyroid_extraction.py` - no qual é feito o carregamento, o pré-processamento e a extração de

¹ Guia de Doenças e Sintomas - Hipertireoidismo. Disponível em:

<https://www.einstein.br/Pages/Doenca.aspx?eid=146>. Acesso em: 29 abr 2022.

² Manual MSD - Hipertireoidismo. Disponível em:

<https://www.msmanuals.com/pt-br/casa/dist%C3%B3rbios-hormonais-e-metab%C3%B3licos/dist%C3%B3rbios-da-tireoide/hipertireoidismo>. Acesso em: 29 abr. 2022.

- características das imagens do *dataset*;
- `thyroid_model.py` - no qual é feita a classificação das imagens pelo modelo de aprendizagem de máquina escolhido.

PRÉ-PROCESSAMENTO DAS IMAGENS E EXTRAÇÃO DE CARACTERÍSTICAS

Após a leitura das imagens cruas dos dois *datasets*, foram feitas algumas manipulações - usando funções do `opencv` -, com o objetivo de melhorar a qualidade das imagens para extração das características:

- `blur` com kernel 3 x 3, para remoção de ruído;
- aumento de contraste com `alpha` 1,5 e `beta` 0, também visando remover ruídos e facilitar a segmentação da tireoide;
- `erosão` com kernel 3 x 3, para segmentar melhor as imagens.

A seguir, cada imagem pré-processada foi convertida em um *array* `numpy` e foi aplicada uma **Análise de Componentes Principais (PCA)** para reduzir os valores dos pixels a um vetor com apenas 6 características para cada imagem.

Em sistemas de diagnóstico assistido por computador ou sistemas CAD (*Computer Aided Diagnostic*), é habitual solicitar aos médicos - antes do processamento das imagens - que selecionem as regiões de interesse ou ROIs (*regions of interest*). Essa etapa ajuda eliminar elementos que não fazem diferença para o diagnóstico e que atrapalham o processamento e a classificação das imagens. No caso de exames de cintilografia de tireoide, por vezes as glândulas salivares aparecem destacadas nas imagens por também serem sensíveis ao contraste. Isso pode dificultar a extração de características e piorar os resultados da classificação das doenças ligadas à tireoide. Tendo esse cenário em mente, foi realizada uma nova etapa de pré-processamento das imagens cruas de ambos os *datasets*:

- definição manual das ROIs de cada imagem;
- `crop` das imagens e manutenção apenas das ROIs;
- filtro de `blur` Gaussiano com kernel 5 x 5, para remoção de ruído;
- ajuste de brilho e de contraste com `alpha` 1,8 e `beta` 8 das imagens filtradas, para destacar melhor os pontos da tireoide com alta atividade metabólica;
- aplicação do filtro *top-hat* com kernel 3x3 nas das imagens filtradas - porém sem ajuste de brilho e contraste. Essa transformação permite realçar objetos brilhantes de interesse em um fundo escuro e, portanto, pode ser útil na identificação de padrões em imagens de cintilografia;
- binarização das imagens já com brilho e contraste ajustados, usando `threshold` `otsu` para definir o melhor limiar para binarização de cada imagem. As imagens foram binarizadas com o objetivo de facilitar a segmentação da tireoide para, posteriormente, se contar o número de pixels correspondentes a cor branca com o objetivo de ter uma métrica de tamanho da tireoide;
- `erosão` das imagens binarizadas com kernel em formato de cruz de 3 x 3 e 2 iterações como parâmetros. Novamente, o objetivo foi facilitar a segmentação da tireoide e, posteriormente, contar o número de elementos presentes em cada imagem. As imagens de doença de Graves tendem a apresentar um único elemento, a tireoide bem definida no formato de “borboleta”, pois todas as áreas da glândula estão hiperativas.

As imagens cortadas de acordo com as ROIs e tratadas foram usadas para extração de outras 46 características:

- **momentos de Hu**, conhecidos também como momentos invariantes - muito utilizados na identificação de padrões em imagens. Os momentos de Hu servem para identificar um objeto, mesmo que tenha sofrido mudança de tamanho ou seja rotacionado. Na captação de imagens médicas, podem haver diferenças significativas de posicionamento do paciente,

distância e enquadramento - além das diferenças anatômicas que naturalmente existem entre os indivíduos. Portanto, os momentos de Hu são indicados como descritores de imagens dessa natureza. Foram extraídos momentos de HU das imagens sem qualquer filtro aplicado (7) e após a aplicação do filtro *top-hat* (7);

- **Local Binary Patterns (LBP)**, que se baseia no fato de que determinados padrões binários locais à região de vizinhança de um pixel são propriedades fundamentais da textura de uma imagem. O histograma de ocorrência destas características é considerado um bom descritor de textura. As características usando LBP (26) foram extraídas das imagens após aplicação do filtro de *blur* Gaussiano;
- **número de pixels brancos** (1), extraídos do histograma das imagens binarizadas;
- **média, desvio padrão e mediana** (3) dos valores dos pixels das imagens normalizadas após ajuste de brilho e contraste;
- **número de objetos por imagem**. Tanto nas imagens binarizadas (1) quanto nas imagens erodidas após a binarização (1), foi encontrado o contorno dos objetos e contado o número de objetos, resultando em duas novas características por imagem.

As 52 características extraídas de cada imagem foram gravadas em um arquivo no formato SVM Light (0 0:0 1:0 2:1 3:1 ... n:0). O primeiro caractere indica o rótulo da classe. A sequência i:v indica o índice da característica e seu valor.

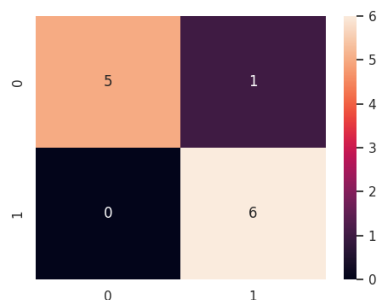
MODELO ESCOLHIDO PARA A CLASSIFICAÇÃO

A escolha do classificador foi feita de modo empírico. Foram testados inicialmente classificadores lineares para descobrir se o problema era linearmente separável: k-NN foi implementado com k igual 3, distância euclidiana e a versão automática do algoritmo do scikit-learn; versão Gaussiana Naïve Bayes, LDA e Perceptron também com os parâmetros padrão das funções do scikit-learn; e, por fim, foi definido como número máximo de iterações para a Regressão Logística 1.000 com os demais parâmetros padrão da função do scikit-learn. Também foi testado o SVM tanto na versão linear quanto na versão não-linear, Radial Basis Function (RBF).

Todos tiveram desempenho inferior ao do classificador escolhido, o **Gradient Boosting**, na implementação do scikit-learn, com os seguintes parâmetros: `n_estimators = 10`, `learning_rate = 1`, `max_depth = 1` e `random_state = 0`. O Gradiente Boosting é um classificador múltiplo *ensemble*, que combina um conjunto de classificadores gerados automaticamente de alguma forma visando diminuir o *bias* (a distância do resultado do classificador do seu objetivo). O desempenho do Boosting pode ser melhorado ao se dar pesos para os classificadores e, no caso do Gradient Boosting, os pesos são encontrados por meio do algoritmo Gradient Descent.

APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Utilizando a implementação apresentada, o algoritmo erra apenas a classificação de 1 das 12 imagens: a imagem "D405923.dcm", pertencente ao paciente ID1, que é diagnosticado com BMT, é classificada equivocadamente como doença de Graves. Com esses **11 acertos**, foi atingida uma **acurácia de 0,916**. A **matriz de confusão** é apresentada abaixo, sendo a classe 0 BMT e a 1 Graves:



A **sensibilidade (recall)** avalia a capacidade do modelo de encontrar todos os pacientes com uma

determinada condição, ainda que alguns pacientes sem esta condição sejam classificados equivocadamente como doentes (falsos positivos). A sensibilidade é dada por $\frac{\text{Verdadeiros positivos}}{(\text{Verdadeiros positivos} + \text{Falsos negativos})}$. O modelo deve ter uma alta sensibilidade quando os falsos negativos são considerados mais prejudiciais que os falsos positivos.

Neste trabalho, o algoritmo atingiu **sensibilidade de 0,83 para a classe BMT** e de **1,00 para a classe Graves**. Isso significa que ele pode ser mais sensível para identificar pacientes com diagnóstico de síndrome de Graves do que aqueles com BMT e que pode confundir o diagnóstico dos pacientes com BMT, classificando equivocadamente o exame como doença de Graves. **No caso da doença de Graves, a alta sensibilidade não é prejudicial. A doença tem outros sintomas clínicos bem característicos além do bócio, como proptose ocular, que podem servir como sinal diferencial para avaliar os falsos positivos.**

Já a **especificidade (precisão)** avalia a capacidade do modelo de encontrar resultados negativos: quem não pertence àquela classe, ainda que classifique equivocadamente alguns pacientes saudáveis como doentes (falsos negativos). A especificidade é dada por $\frac{\text{Verdadeiros negativos}}{(\text{Verdadeiros negativos} + \text{Falsos positivos})}$. Busca-se uma alta especificidade quando os falsos positivos são considerados mais prejudiciais que os falsos negativos. O algoritmo proposto atingiu **especificidade de 1,00 para a classe BMT** e de **0,86 para a classe Graves**.

Por fim, o **F1-score** é uma média harmônica das duas métricas: sensibilidade e especificidade. Um valor baixo de F1-score indica que uma das duas métricas está baixa. O modelo atingiu **F1-score de 0,91 para a classe BMT** e de **0,92 para a classe Graves**.

Nota-se abaixo que, na **imagem da classe BMT classificada de forma errada** (3ª da 1ª linha, da esquerda para direita) pelo modelo, a tireoide tem um **padrão visual muito similar às tiroides das imagens dos pacientes com doença de Graves**. Dadas as limitações da base (*dataset* pequeno e imagens de Medicina Nuclear, cuja qualidade de resolução é menor), **consideramos que o modelo obteve um bom desempenho** ao classificar imagens de pacientes com BMT e doença de Graves.

