

TAREFA 1 - PROCESSAMENTO DE IMAGEM

Nome: Arthur Lodetti Gonçalves; Gustavo Piacentini da Silva.

Matéria: Processamento de Imagem

Professor: Gilmário Barbosa dos Santos

Data: 09/10/2025

O projeto pode ser acessado no GitHub, através deste link:

<https://github.com/GroudOn19/PIM>

Objetivo do trabalho:

O objetivo principal deste trabalho é aplicar os conceitos vistos em sala de aula na área de segmentação de volume, operações preparatórias como rotulação e identificação de grupos conexos. A principal tarefa é realizar este tipo de algoritmo, sem utilizar bibliotecas prontas, para a realização destas operações e analisar um volume tomográfico 3D de uma neoplasia utilizando a linguagem Python.

Caracterização das células:

- Proliferativas: são as células doentes capazes de se reproduzir;
- Quiescentes: células vivas mas inativas, são capazes de se ativar na forma de proliferativas ou de morrer, neste caso elas se tornam células necróticas;
- Necróticas: são células mortas.

O trabalho consistiu em analisar dois exemplos de vizinhança 3D: conectividade-6 e conectividade-26. Foram abordados as seguintes extrações de dados para cada tipo de célula: número total de células, histogramas contendo o volume de cada agrupamento e uma visualização 3D da morfologia dos componentes celulares. Compreender a diferença entre os dois tipos de conectividades é uma chave essencial para o estudo.

Metodologia Utilizada:

O processo pode ser aplicado independente dos dados iniciais, contanto que seja fornecida uma matriz numpy 3D preenchida com os valores de cinza corretos, o algoritmo pode segmentar a tomografia e trazer dados relevantes à análise. Ele segue o fluxo exibido no diagrama abaixo:



Extração de Dados

Simplesmente utilizamos a biblioteca pickle e a função `pickle.load()` para extrair os dados do arquivo `volume_TAC` para uma matriz `numpy.ndarray`.

Segmentação

Dentro do loop principal do algoritmo, a função `rotula_tomografia_3d()` é utilizada 3 vezes, uma para cada tipo de célula presente na tomografia. Essa função recebe três parâmetros, o volume, que são os dados extraídos do arquivo, o `tipo_cinza`, que é o tom de cinza que representa a célula a ser analisada, e a conectividade, que pode ser 6 ou 26. A diferença entre as conectividades é o objeto principal de análise desse trabalho.

O processo funciona da seguinte maneira:

- Cria uma matriz de mesma dimensão da original, e zera cada um de seus voxels. Esse será o `volume_rotulado`.
- Dependendo da conectividade, seleciona uma lista de vizinhos, que são movimentos ao redor do voxel em análise, para representar os 6 ou 26 vizinhos dele
- Percorre cada voxel do volume sistematicamente, verificando duas condições:
 - Se o voxel tem o mesmo valor que o `tipo_cinza`;
 - Se o voxel ainda tem valor 0 dentro do `volume_rotulado`.
- Se as duas forem verdade, temos um novo agrupamento da célula em análise
- A partir dela, inicia uma BFS verificando todos os membros de sua vizinhança fazendo os mesmos testes, e rotulando-os caso verdadeiro.
- Ao final do processo temos um dicionário de dados com todos os agrupamentos enumerados, e seus tamanhos (quantidade de voxels).
- E, apesar de não utilizarmos neste trabalho, o algoritmo retorna o `volume_rotulado`, com todos os agrupamentos de um tipo de células rotuladas.

Como o volume possui uma grande quantidade de voxels, o número de comparações cresce rapidamente, mas falaremos melhor disso na resposta à questão 4.3.

Processamento

A função `conta_celula_por_tipo()` retorna as quantidades de cada célula dentro do volume. Junto a isso, para cada tipo de célula separamos o maior e menor agrupamento, e calculamos a média de voxels por agrupamento.

Além disso, o tempo gasto para segmentar o volume é capturado usando a biblioteca time e registrando o tempo de início e fim das operações, subtraindo-os e somando o resultado ao tempo total. Ao fazer isso para as três execuções da função `rotula_tomografia_3d()`, podemos comparar o tempo total gasto para segmentação entre as conectividades 6 e 26.

Por final, a quantidade de comparações que acontecem dentro da função de rotulação também é registrada para posterior comparação.

Plotagem

Dois tipos de gráficos são formados a partir dos resultados. Um deles é uma projeção 3D do volume rotulado, uma para cada tipo de célula, sendo as proliferativas vermelhas, as quiescentes azuis e as necróticas pretas. Junto a esse gráfico, exibimos o total de voxels daquele tipo celular.

O outro tipo é um histograma, exibindo no eixo Y o número de voxels, e no eixo X uma barra para cada agrupamento. O eixo vertical desses gráficos possui uma escala logarítmica devido à alta probabilidade de que um dos grupos possua uma quantidade de voxels significativamente maior que o resto. Junto a eles, exibimos os números relacionados aos agrupamentos, maior número de voxels, menor número de voxels e média por grupo.

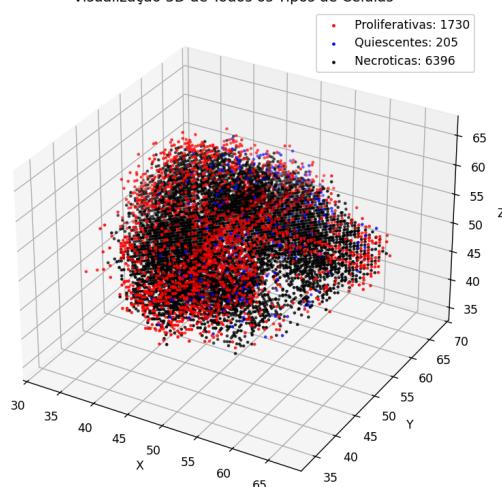
Por último, um gráfico 3D com todos os voxels do volume, coloridos de acordo com seu tipo de célula, é exibido na tela, com os totais de cada um dos tipos.

Resultados Obtidos:

Saída do terminal após executar o código “[main.py](#)”:

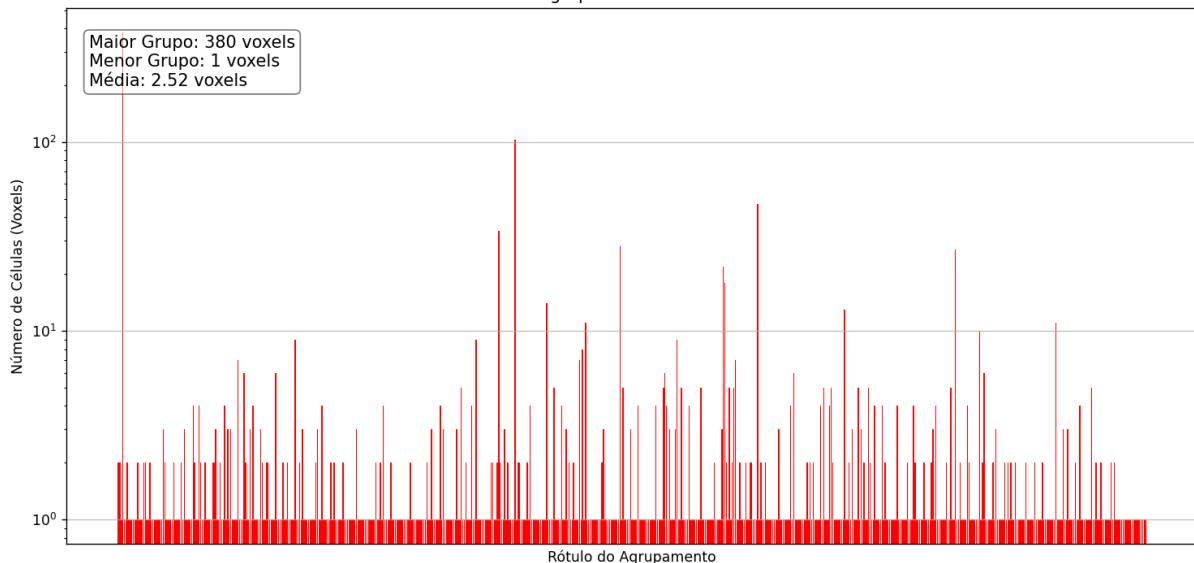
Conectividade-6:

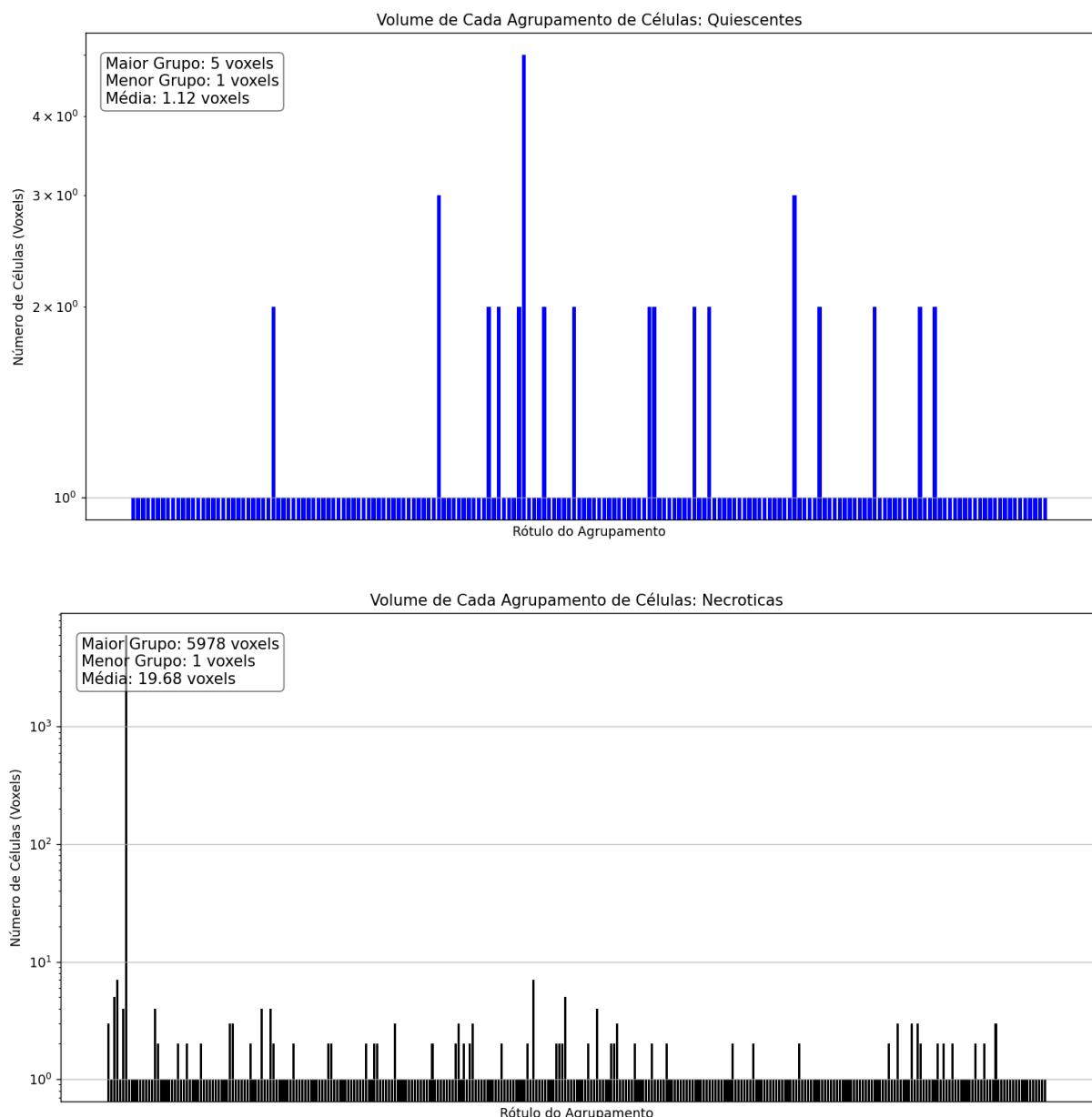
Visualização 3D de Todos os Tipos de Células



Histogramas de Distribuição:

Volume de Cada Agrupamento de Células: Proliferativas





Levantamento de dados para cada tipo de célula:

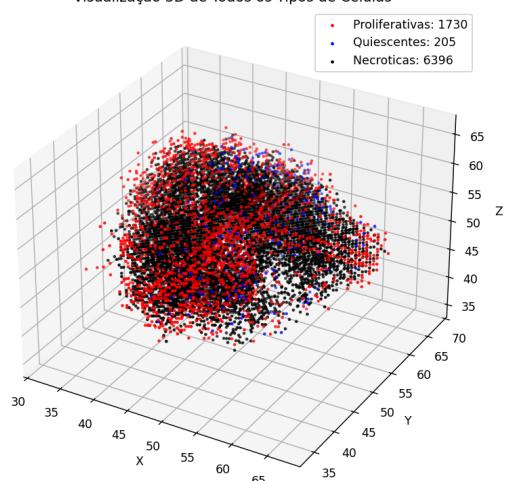
- Células Proliferativas (255): Foram encontrados 687 componentes, com uma média de 2.52 voxels por agrupamento. O maior grupo contém 380 voxels.
- Células Quiescentes (200): Foram encontrados 183 componentes, com uma média de 1.12 voxels por agrupamento. O maior grupo contém 5 voxels.
- Células Necróticas (140): Foram encontrados 325 componentes, com uma média de 19.68 voxels por agrupamento. O maior grupo contém 5978 voxels.

Tabela com organização dos dados:

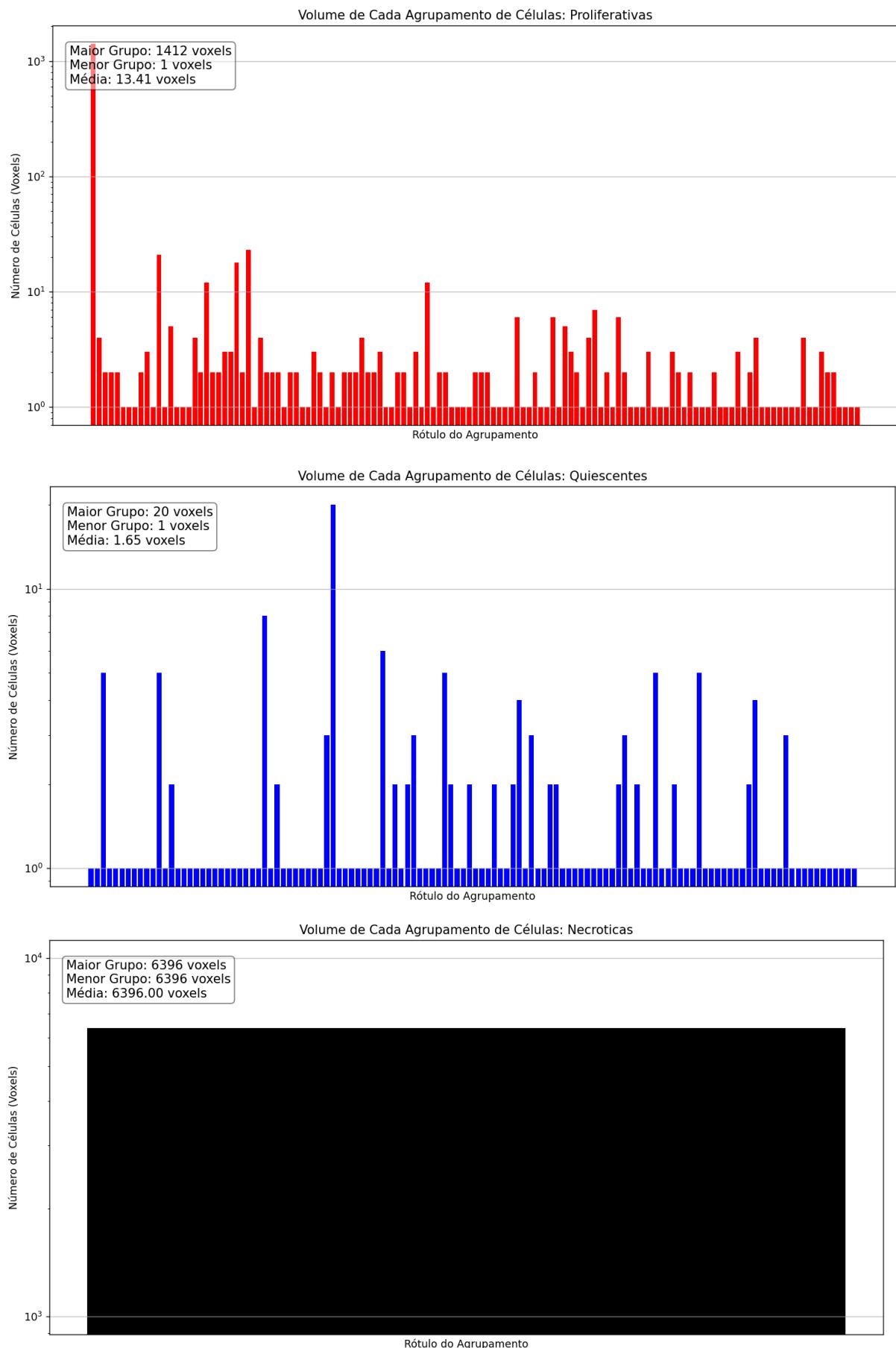
Conectividade-6	Necróticas(140)	Quiescentes(200)	Proliferativas(255)
Nº de componentes	325	183	687
Quantidade de voxels	6396	205	1730
Tamanho médio	19,68	1.12	2.52
Tamanho máximo	5978	5	380

Conectividade-26:

Visualização 3D de Todos os Tipos de Células



Histogramas de distribuição:



Levantamento de dados para cada tipo de célula:

- Células Proliferativas (255): Foram encontrados 129 componentes, com uma média de 13.41 voxels por agrupamento. O maior grupo contém 1412 voxels. A redução de número de componentes acrescida com um número de voxels elevado no maior grupo evidencia uma análise mais criteriosa e coesa com a conectividade-26.
- Células Quiescentes (200): Foram encontrados 124 componentes, com uma média de 1.65 voxels por agrupamento. O maior grupo contém 20 voxels. Os agrupamentos das células não tiveram tanta mudança, evidenciando que é uma distribuição mais isolada.
- Células Necróticas (140): Foram encontrados 1 componente, com uma média de 6396 voxels por agrupamento. O maior grupo contém 6396 voxels. Isso nos evidencia a maior mudança entre as conectividades, visto que o número de componentes reduziu para 1, o que antes era 325.

Tabela com organização dos dados:

Conektividade-26	Necróticas(140)	Quiescentes(200)	Proliferativas(255)
Nº de componentes	1	124	129
Quantidade de voxels	6396	205	1730
Tamanho médio	6396	1.65	13.41
Tamanho máximo	6396	20	1412

Comparativo entre conectividade-6 com conectividade-26:

Após esse levantamento de dados, conseguimos observar a grande diferença entre a aplicação utilizando conectividade-6 e conectividade-26. Inicialmente, observamos uma estrutura altamente fragmentada com 1125 componentes. No entanto, com a conectividade-26 há uma queda na quantidade para 254 componentes distintos. Essa grande diferença entre as análises não sugere erro algum, afinal a conectividade-26 é mais abrangente por conectar voxels que se tocam até nas diagonais, agrupando diversos componentes que a conectividade-6 via como separados. Com isso, compreendemos a estrutura morfológica do tumor além de afirmar a fidelidade da conectividade-26 com a realidade, sendo essencial para uma análise mais precisa.

4.2. O que aconteceria se fosse aplicada a conectividade-4 ou a conectividade-8 na segmentação do volume na questão?

A abordagem desses tipos de conectividade bidimensionais seria inadequada para a análise de um volume 3D. Esta aplicação trataria o volume como diversas “fatias” 2D. Isso nos evidencia um grande erro, pois cada plano XY seria tratado de forma isolada, voxels adjacentes não seriam reconhecidos como vizinhos, ignorando a conectividade pelo eixo Z. Com isso, células que atravessam o plano Z, perdem esse “valor” e seriam fragmentadas em diversas células separadas em cada plano, levando a um estudo incorreto a respeito do tumor.

4.3. A respeito da escolha do tipo de conectividade em 3D (Figura 3), discuta o impacto sobre o custo computacional e sobre a acurácia do resultado da segmentação.

Por analisar um número maior de vizinhos por voxel, a conectividade-26 instintivamente tem um custo computacional maior. Entretanto, para validar essa suspeita, pensamos em duas maneiras de analisar a complexidade dos algoritmos, um deles sendo mais confiável e replicável que o outro.

Complexidade Temporal

Para análise de complexidade temporal, marcamos o tempo de início e fim de cada segmentação, e somamos a diferença entre os dois no total, isso para cada tipo de célula

Executando o algoritmo 10 vezes seguidas, a média da diferença de tempo entre a conectividade-6 e a conectividade-26 foi de 0,0409 segundos. A maioria das execuções com a primeira ficavam na casa dos 3,3 segundos, enquanto as da segunda terminavam na casa dos 3,7 segundos, representando um aumento de aproximadamente 12% de uma para a outra.

Entretanto, esse método de análise é vulnerável a muitos fatores externos, desde a velocidade de processamento da máquina, até a quantidade de processos competindo pelos mesmos recursos da CPU. Portanto não deve ser interpretado como fato.

Número de Comparações

Dentro da função de segmentação, um contador de comparações foi colocado. Os resultados nos deixaram confusos por um tempo pois, para a conectividade-6 tivemos 6.249.930 comparações e para conectividade-26 obtemos 7.083.030 delas.

Mas analisando um pouco a fundo, descobrimos que isso se deve à natureza de nosso algoritmo. O volume tem dimensões 100x100x100, resultando em um total de 1.000.000 de voxels. Para cada um deles fizemos as duas comparações mostradas na metodologia, e na execução do loop principal, o algoritmo é executado 3 vezes, uma para cada tipo de célula, ou seja, o número mínimo de comparações era 6.000.000.

A verdadeira análise se revela quando reduzimos esse número dos resultados obtidos, restando 249.930 comparações para a conectividade-6 e 1.083.030 comparações para a

conectividade-26. Isso revela um aumento de aproximadamente 333% no número de comparações entre os dois métodos. Ou seja, um claro indicador do custo do segundo método.

Conclusão:

A implementação e avaliação do algoritmo de segmentação volumétrica e rotulação de componentes conexos permitiu compreender de forma prática a comparação e o funcionamento entre a conectividade-6 e conectividade-26. A partir das diferentes técnicas de agrupamento, é possível definir diferenças na identificação e quantidade de grupos, conduzindo a uma interpretação morfológica singular de cada conectividade. Por mais que seja menos custosa, a conectividade-6 nos mostrou uma segmentação mais fragmentada, com 1125 componentes, enquanto a outra retornou 254 componentes conexos. Essa diferença de abordagem evidencia a credibilidade da conectividade-26 para casos de análise e distribuição volumétrica. Os resultados obtidos no trabalho mostram a importância da escolha adequada de conectividade conforme o objetivo da análise. Por fim, esse desenvolvimento serviu como aprendizado e aplicação dos conceitos vistos em sala de aula, a respeito de segmentação volumétrica e rotulação.